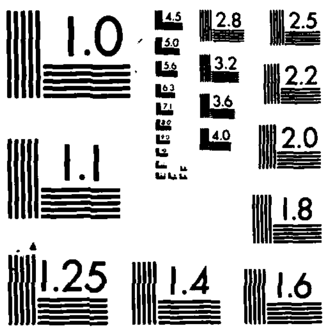
AD-A159 547 REPEATED SIGNIFICANCE TESTS FOR EXPONENTIAL FAMILIES 1/1
(U) STANFORD UNIV CA DEPT OF STATISTICS I HU AUG 85
TR-34 N00014-77-C-0306

UNCLASSIFIED F/G 12/1 NL

END
FILMED
DTIC

END
FILMED
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

1

AD-A159 547

REPEATED SIGNIFICANCE TESTS FOR EXPONENTIAL FAMILIES

by

Inchi Hu
Stanford University

TECHNICAL REPORT NO. 34
AUGUST 1985

PREPARED UNDER CONTRACT
N00014-77-C-0306 (NR-042-373)
FOR THE OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

REPEATED SIGNIFICANCE TESTS FOR
EXPONENTIAL FAMILIES

by

Inchi Hu
Stanford University

TECHNICAL REPORT NO. 34
AUGUST 1985

PREPARED UNDER CONTRACT
N00014-77-C-0306 (NR-042-373)
FOR THE OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

Chapter 1

Introduction

In this thesis we study the significance levels and powers of repeated significance test (RST). A RST is the sequential version of the generalized likelihood ratio test. To be more precise, let us consider the following testing problems Let $X_1, X_2, \cdots$ be i.i.d. according to the distribution function F_θ where $\{F_\theta\}$ form a multiparameter exponential family. By that we mean F_θ has the form $F_\theta(dx) = e^{\theta' x - \psi(\theta)} F_0(dx)$ for some smooth function $\psi(\cdot)$ from the parameter space Θ into R' and some distribution function F_0 over R^d. It is well known that $E_\theta X_1 = \mu(\theta) = \nabla\psi(\theta)$. Moreover, there is no loss of generality to assume that $\mu(0) = 0$. Sometimes it is convenient to index this family by μ and write F_μ. Let Θ_0 be a proper subset of $\Theta \subset R^d$. If we want to test $H_0 : \theta \in \Theta_0$ against $H_1 : \theta \notin \Theta_0$. The generalized log likelihood ratio statistic after observing $X_1, X_2, \cdots, X_n$ for this testing problem is

$$n\Lambda(S_n/n) = \sup_{\theta\in\Theta} \ell_n(\theta) - \sup_{\theta\in\Theta_0} \ell_n(\theta) = n\phi(S_n/n) - n\phi_0(S_n/n)$$

where $\ell_n(\theta) = \theta' S_n - n\psi(\theta)$ is the log likelihood after observing $X_1, \cdots, X_n$, $S_n = \sum_{i=1}^n X_i$ and

$$\phi(x) = \sup_{\theta\in\Theta}[\theta' x - \psi(\theta)], \quad \phi_0(x) = \sup_{\theta\in\Theta_0}[\theta' x - \psi(\theta)].$$

The RST is defined in terms of the following stopping rule

$$T = \inf\{n \geq m_0,\ n\Lambda(S_n/n) > a\}.$$

It stops sampling at $T \wedge m$ and rejects H_0 when $T \leq m$. The significance level and power of the RST are given by

$$\max_{\Theta_0} P_\theta\{T \leq m\}$$

and $P_\theta\{T \leq m\}$ $\theta \notin \Theta_0$ where P_θ denotes the probability law under which $X_1, X_2, \cdots$ are i.i.d. according to the distribution function F_θ. In some cases when one expects a small deviation from the null hypothesis and wants to increase the power, one may use a modified version of the RST. The MRST rejects H_0 when either $T \leq m$ or $T > m$ and $m\Lambda(S_m/m) > c$ for some $c < a$.

Observe that when we fix c and let a tend to ∞ then it is unlikely that the log likelihood ratio process $n\Lambda(S_n/n)$ will cross the level a before time m and the rejection region of the corresponding MRST reduces essentially to $\{m\Lambda(S_m/m) > c\}$ which is exactly the rejection region of a fixed sample test. On the other hand if we set $a = c$ then the corresponding MRST is just RST. So the MRST can be thought of as a family of tests interpolating the fixed sample size test and the RST.

Underlying this interpolation, there is a trade-off between expected sample size and power, that is, as a moves from c to ∞ the power of MRST increases to that of a fixed sample test at the cost of increasing the expected sample size. So with MRST at hand the designer of an experiment has one more degree of freedom to choose in fulfilling his needs. If he thinks power is more important he may choose a MRST with a substantially larger than c. If smaller expected sample size is desired he may choose a close to c.

The power of the MRST is given by

$$\begin{aligned} &P_\theta\{T < m\} + P_\theta\{T \geq m,\ m\Lambda(S_m/m) > c\} \\ &\qquad = P_\theta\{m\Lambda(S_m/m) > c\} + P_\theta\{T < m,\ m\Lambda(S_m/m) \leq c\} \end{aligned} \tag{1}$$

The quantity (1) also appears on other occasions. Siegmund (1985) suggests defining the attained significance levels for a RST as follows:

(i) If $T = m_0$ and $m_0\Lambda(S_{m_0}/m_0) = z > a$ then the attained level is $\sup_{\Theta_0} P_\theta\{m_0\Lambda(S_{m_0}/m_0) > z\}$.

(ii) If $T = n \in (m_0, m]$ the attained level is $\sup_{\Theta_0} P\{T \leq n\}$.

(iii) If $T > m$ and $m\Lambda(S_m/m) = c$ then the attained level is

$$\begin{aligned}&\sup_{\Theta_0}[P_\theta\{T \leq m\} + P_\theta\{T > m,\ m\Lambda(S_m/m) \geq c\}] \\ &\quad = \sup_{\Theta_0}[P_\theta\{m\Lambda(S_m/m) > c\} + P_\theta\{T \leq m,\ m\Lambda(S_m/m) \leq c\}].\end{aligned}$$

In case (iii) above the attained significance level is of the same form as (1).

In this thesis we only consider a special kind of Θ_0, e.g.

$$\Theta_0 = \{\theta : \theta_1 = \cdots = \theta_{d_1} = 0\} \quad d_1 \leq d$$

By reparametrization Θ_0 can be generalized to $\Theta_0' = \{A\theta : \theta \in \Theta\}$ where A is a $d \times d$ matrix. Typically the significance level and power of the RST and MRST cannot be computed exactly and some sort of approximation is required. Approximations for significance levels of the RST in exponential families have been provided by Woodroofe (1978) and Lalley (1983). Their setting is more general than that above, but their methods are not as successful in approximating the power of the RST and to power and significance level of the MRST. In what follows we shall exhibit with a simple example three methods which have been developed by previous authors. Let $X_1, X_2, \cdots$ be i.i.d. according to $N(\theta, 1)$. We want to test $\theta = 0$ against $\theta \neq 0$. The RST in this case is defined by the stopping rule

$$T = \inf\left\{n \geq m_0,\ S_n^2/(2n) > a\right\}.$$

1.1. Forward Method.

The essential ingredients of this method are the likelihood ratio of a mixture measure Q and the probability measure P_0 under the null hypothesis, and the Wald likelihood ratio identity. Let $Q(A) = \int_{-\infty}^{\infty} P_\theta(A)d\theta$ then

$$\frac{dQ}{dP_0}(S_1, \cdots, S_n) = \int_{-\infty}^{\infty} \exp(\theta S_n - n\theta^2/2n)d\theta = (2\pi/n)^{1/2}e^{S_n^2/2n}.$$

Here the notation $\frac{d\mu}{d\nu}(Y)$ means that μ and ν are considered to be measures on the σ-field generated by Y, and $\frac{d\mu}{d\nu}(Y)$ is the Radon-Nikodym derivative of the restricted measures.

By Wald's likelihood ratio identity

$$\begin{aligned} P_0\{T \le m\} &= E_Q\{(T/2\pi)^{1/2} e^{-S_T^2/2T};\ T \le m\} \\ &= \int_{-\infty}^{\infty} E_\theta\{(T/2\pi)^{1/2} \exp[-(a + R_m(T))];\ T \le m\} d\theta \\ &= (a/\pi)^{1/2} e^{-a} \int_{-\infty}^{\infty} E_\theta \left\{(T/2a)^{1/2} e^{-R_m(T)};\ T \le m\right\} d\theta \end{aligned}$$

where $R_m(T) = (S_T^2/2T - a)$ is the corresponding excess over the boundary for this problem.

Before we go any further, let me introduce some notation. Throughout this work I use $R(T)$ to denote the excess over the boundary corresponding to the stopping time T. Usually the stopping time depends on a scale parameter m. To emphasize the dependence on m I also write $R_m(T)$ or R_m if it does not cause confusion, and $R_\infty(T)$ or R_∞ the corresponding limit in distribution as $m \to \infty$.

If a, m, m_0 tend to ∞ in such a way that $(2a/m)^{1/2} = \theta_1 < \theta_0 = (2a/m_0)^{1/2}$ then an argument using the strong law of large numbers shows that with P_θ-probability one

$$(T/2a)^{1/2} 1_{(m_0 < T \le m)} \to \theta^{-1}\, 1_{[\theta_1, \theta_0]}(\theta)$$

and

$$P_0\{T \le m\} \sim (a/\pi)^{1/2} e^{-a} \int_{\theta_1}^{\theta_0} \theta^{-1} E_\theta\{e^{-R_\infty(T)}\} d\theta$$

. Now $E_\theta(e^{-R_\infty(T)})$ can be approximated using nonlinear renewal theory developed by Lai and Siegmund (1977, 1979) or Woodroofe (1976a) and the approximation is completed.

This method has been generalized to RSTs for curved exponential families by Lalley (1983).

1.2. The Backward Method.

The backward method which is due to Siegmund (1985) sets its primary goal on approximating the conditional probability $P_\xi^{(m)}(A) = P\{A \mid S_m = \xi\}$, which by sufficiency of S_m is independent of θ. Then the power and significance level may be obtained by unconditioning with respect to the distribution of S_m. In Chapter 3,4 we shall generalize this method to multiparameter exponential families.

The essence of this method involves randomizing the starting point of a process, then treating it like a process running backward from the point of conditioning. Let $P_{\lambda,\xi}^{(m)}(A) = P(A \mid S_0 = \lambda,\ S_m = \xi)$ and $T^* = \sup\{n \le m,\ S_n^2/2n > a\}$. Observe that $P_{0,\xi}^{(m)}(T \le m) = P_{0,\xi}^{(m)}(T^* \ge m_0)$. Let

$$\tilde{P}_{\xi}^{(m)}(A) = \int_{-\infty}^{\infty} P_{\lambda,\xi}^{(m)}(A)(2\pi m)^{-1/2}\exp\{-[(\lambda-\xi)^2/2m]\}d\lambda$$

Then

$$\frac{d\tilde{P}_{\xi}^{(m)}}{dP_{0,\xi}^{(m)}}(S_n,\cdots,S_m) = \left(\frac{n}{m}\right)^{1/2}\exp\left(\frac{S_n^2}{2n} - \frac{\xi^2}{2m}\right).$$

Since under the reversed time scale T^* is a stopping time, Wald's likelihood ratio identity gives

$$P_{\xi}^{(m)}(T^* \ge m_0) = \tilde{E}_{\xi}^{(m)}\left\{\left(\frac{m}{T^*}\right)^{1/2}\exp\left(\frac{\xi^2}{2m} - \frac{S_{T^*}^2}{2T^*}\right);\ T^* \ge m_0\right\}.$$

The $P_{\xi}^{(m)}$ distribution of S_n, $n = m, m-1, \cdots$ running backward from $S_m = \xi$ is the same as the P_0 distribution of $\xi - S_n$, $n = 0, 1, \cdots$ running forward.

Hence the expectation above equals

$$E_0\left\{\left(\frac{m}{m-\tau}\right)^{1/2}\exp\left[\frac{\xi^2}{2m} - \frac{(S_\tau+\xi)^2}{2(m-\tau)}\right];\ \tau \le m - m_0\right\}$$

where $\tau = \inf\{n \ge 1,\ (\xi + S_n)^2/[2(m-n)] > a\}$. Assume that $\theta_1 = (2a/m_1)^{1/2}$ and $\xi_0 = m^{-1}\xi$. A law of large numbers argument shows that $\tau/m \to 1-(\xi_0/\theta_1)^2$ with probability one as $m \to \infty$. The quantity above is approximated by

$$(m\theta_1/\xi)\exp(-a + \xi^2/2m)E_0\{e^{-R_\infty(\tau)}\}$$

where

$$R_m(\tau) = \{(S_\tau - \xi)^2/[2(m-\tau)]\} - a$$

is the excess over the boundary at the stopping time τ. Again nonlinear renewal theory can be used to obtain the asymptotic distribution of $R_m(\tau)$ and the approximation for $P_{\xi}^{(m)}(T \le m)$ is completed. Unconditioning ξ using the marginal distribution of S_n under P_0 yields $P_0(T \le m)$.

We may uncondition ξ using P_θ with $\theta \neq 0$ and obtain an approximation to the powers. Unfortunately the result is not a bona fide asymptotic expression, although the numerical results show that it is a very good approximation. See Siegmund (1985) Section 9.3 for details.

1.3. Woodroofe's Method.

This method, which was developed by Woodroofe, is quite different from the two methods described above. It does not use Wald's likelihood ratio identity. The method first approximates $P_0\{T = n\}$ then estimates $P_0\{T \leq m\}$ by summation. Observe that $P_0\{T = m\} \sim 2P\{T_+ = n\}$ where $T_+ = \inf\{n \geq m_0,\ S_n > \sqrt{2na}\}$

$$P_0\{T_+ = n\} = \int_{\sqrt{2an}}^{\infty} P_\xi^{(n)}(T_+ > n-1)(2\pi n)^{-1/2} \cdot \exp(-\xi^2/2n)d\xi.$$

It is easy to see that the only values of ξ which are of first order importance are $\sqrt{2an} + 0(1)$. In this range we can approximate the curve $\sqrt{2na}$ by its tangent and the conditional random walk by an unconditional one (with drift ξ_0). That is, let $\xi = \sqrt{2an} + y$ where y is arbitrary but fixed.

$$\begin{aligned} P_\xi^{(n)}\{T_+ > n-1\} &= P_0\{S_k < \sqrt{2ak} \text{ for all } m_0 \leq k \leq n-1 \mid S_n = \sqrt{2an} + y\} \\ &= P_0\{S_n - S_k > y + \sqrt{2a}(n^{1/2} - k^{1/2}) \text{ for all } m_0 \leq k < n \mid S_n = \sqrt{2an} + y\} \\ &= P_0\{S_i > y + \sqrt{2a}(n^{1/2} - (n-i)^{1/2}) \text{ for all } 1 \leq i \leq n - m_0 \mid \frac{S_n}{n} = \mu^* + o(1)\}. \end{aligned}$$

Observe that $\sqrt{2n}[n^{1/2} - (n-i)^{1/2}] = \frac{1}{2}\sqrt{2a}n^{-1/2}i + 0(\sqrt{2a}n^{-1/2}n^{-3/2}i^2) \to \frac{1}{2}\mu^* i$ if n and a tend to infinity in such a way that $(2a/n)^{1/2} \to \mu^*$. The conditional probability above is asymptotically equivalent to

$$P_{\mu^*}\{S_i > y + \frac{1}{2}\mu^* i \text{ for all } i \geq 1\} = P_{\mu^*/2}\{S_i > y \text{ for all } i \geq 1\}$$

To continue we need

Lemma 1. (Woodroofe, 1982, p. .) Assume $\mu = EX_1 > 0$. Let $M = \min(S_1, S_2, \cdots)$. Then for $x > 0$

$$[E(S_{\tau_+})]^{-1}P\{S_{\tau_+} > x\} = \mu^{-1}P\{M > x\} \quad \text{where} \quad \tau_+ = \inf\{n \geq 1,\ S_n > 0\}$$

Proof: Define $\alpha = \sup\{n : S_n = M\}$ and $\tau_- = \inf\{n \geq 1,\ S_n \leq 0\}$. Consider the probability $P\{\alpha = n,\ M > x\}$. By restarting the random walk at time n, we find that

$$P\{\alpha = n,\ M > x\} = P\{\tau_- = \infty\} \cdot P\{S_n \leq S_i, \quad i = 1, \cdots, n-1 \quad S_n > x\}.$$

A time reversal argument shows that

$$\begin{aligned} P\{S_n \leq S_i\, i = 1, \cdots, n-1,\ S_n > x\} &= P\{S_i < 0,\ i = 1, \cdots, n-1\ S_n > x\} \\ &= P\{\tau_+ = n,\ S_{\tau_+} > x\}. \end{aligned}$$

Combining all the results above, we get

$$\begin{aligned} P\{M > x\} &= \sum_{n=1}^{\infty} P\{\alpha = n, M > x\} = \sum_{n=1}^{\infty} P\{\tau_- = \infty\} P\{S_n \leq S_i\, i = 1, \cdots, n-1, S_n > x\} \\ &= P\{\tau_- = \infty\} \sum_{n=1}^{\infty} P\{\tau_+ = n, S_{\tau_+} > x\} = (E\tau_+)^{-1} P\{S_{\tau_+} > x\} = \mu (ES_{\tau_+})^{-1} P\{S_{\tau_+} > x\} \end{aligned}$$

We have used the duality relation $P\{\tau_- = \infty\} = (E\tau_+)^{-1}$ and Wald's identity $ES_{\tau_+} = \mu E\tau_+$ in the fourth and fifth equality above. The proof is completed.

By lemma 1 and the argument given above

$$\begin{aligned} P_0\{T = n\} &\sim 2P_0\{T_+ = n\} \\ &\sim 2\int_0^{\infty} \mu^*[E_{\mu^*/2}(S_{\tau_+})]^{-1} P_{\mu^*/2}\{S_{\tau_+} > x\}(2\pi n)^{-1/2} \cdot \exp\{-[(2an)^{1/2} + y^2]/2n\}dy \\ &\sim n^{-1}(a/\pi)^{1/2} e^{-a} \int_0^{\infty} [E_{\mu^*/2}(S_{\tau_+})]^{-1} P_{\mu^*/2}\{S_{\tau_+} > x\} e^{-\mu^* y} dy. \end{aligned}$$

The integral above equals $\lim_{a \to \infty} E_{\mu^*} \exp[-R_a(T)]$ by nonlinear renewal theory, where $R_a(T) = S_T^2/2T - a$. Summing over n and approximating the sum by an integral yields the desired result. Now we are in a position to make brief comments on the three methods described above.

If one were only concerned with the significance levels of the RST then the forward method is the most general of the three. If one wants a second order approximation to the significance level then Woodroofe's method seems to be the appropriate method to use. Since the Monte Carlo results in Chapter 5 show that the "obvious" second order correction works quite well, the complicated second order correction developed by Takahashi and Woodroofe

(1981, 1982) seems to be unnecessary (at least in those cases). But if we restrict our attention to the linear hypothesis, then the backward method produces the most fruitful results, that is, it can be used to approximate the significance level, power, and p-values of the RST and MRST. One of the major contributions of this thesis is to generalize the backward method to multiparameter exponential families.

The rest of the thesis is organized as follows. In Chapter 2 the simple null hypothesis case is considered and a theorem which relates the excess over the boundary by the forward and backward process is proved. Chapter 3 deals with the composite hypothesis problem. The results there indicate the necessity for studying the conditional renewal theory which is the topic of Chapter 4. An application to the "change point" problem is also given. Chapter 5 contains a careful treatment of an important example: the repeated t-test. The numerical approximations of powers and significance levels of RST and MRST and the results of corresponding Monte Carlo experiments are also reported. Finally, the Appendix gives some details of the numerical computation performed in Chapter 5.

Chapter 2

The Simple Null Hypothesis Case

When Θ_0 contains only one point the backward method generalizes to the multiparameter exponential family easily. Without loss of generality we assume $\Theta_0 = \{0\}$. In this case the stopping rule is $T = \inf\left\{n \geq m_0,\ n\phi\left(\frac{S_n}{n}\right) > a\right\}$. Define $T^* = \sup\left\{n \leq m,\ n\phi\left(\frac{S_n}{n}\right) > a\right\}$ then $P\{T \leq m\} = P\{T^* \geq m_0\}$. Now

$$\frac{dP^{(m)}_{\lambda,\xi}}{dP^{(m)}_{0,\xi}}(S_n, \cdots, S_m) = \frac{f_n(S_n - \lambda) f_m(\xi)}{f_n(S_n) f_m(\xi - \lambda)}$$

where $f_n(\cdot)$ is the density of S_n under F_0, we assume f_n exist and satisfy the condition given in Proposition 1 below.

$$P^{(m)}_{\lambda,\xi}(A) = P_0(A \mid S_0 = \lambda,\ S_m = \xi).$$

Also define $Q(A) = \int P^{(m)}_{\lambda,\xi}(A) f_m(\xi - \lambda) d\lambda$. The likelihood ratio of Q with respect to $P^{(m)}_{\xi}$ is easily calculated as

$$L_n \equiv \frac{dQ}{dP^{(m)}_{\xi}}(S_n, \cdots, S_m) = \int \frac{dP^{(m)}_{\lambda,\xi}}{dP^{(m)}_{\xi}}(S_n, \cdots, S_m) f_m(\xi - \lambda) d\lambda = \frac{f_m(\xi)}{f_n(S_n)}.$$

The Q distribution of S_n, $n = m, m-1, \cdots$ running backward from $S_m = m\xi_0$ is the same as the P_0 distribution of $m\xi_0 - S_n$, $n = 0, 1, \cdots$ running forward. Under the reverse time scale T^* is a stopping time so the Wald likelihood ratio identity gives

$$P^{(m)}_{\xi}(T \leq m) = P^{(m)}_{\xi}(T^* \geq m_0) = E_Q\left\{L^{-1}_{T^*};\ T^* \geq m_0\right\} = E_Q\left\{\frac{f_{T^*}(S_{T^*})}{f_m(\xi)};\ T^* \geq m_0\right\}.$$

Let $\tau = \inf\{n \geq 1,\ (m-n)\phi\left(\frac{\xi - S_n}{m-n}\right) > a\}$. Note that under Q $m - \tau$ has the same distribution as T under P_0 so the expectation above equals

$$E_0\left\{\frac{f_{m-\tau}(\xi - S_\tau)}{f_m(\xi)};\ \tau \le m - m_0\right\}.$$

It is not hard to see that if

$$t_{\xi_0} = \inf\left\{t;\ \frac{a_0}{1-t} = \phi\left(\frac{\xi_0}{1-t}\right)\right\} \tag{2.1}$$

exist and $0 < t_0 < 1 - \frac{m_0}{m}$ then $\tau/m \to t_{\xi_0}$ with probability one, where $a = ma_0$, $\xi = m\xi_0$, and the expectation above is approximated by

$$E_0\left\{\frac{f_{m-\tau}(\xi - S_\tau)}{f_m(\xi)}\right\}.$$

To continue the computation we need

Proposition 1. If for some integer n_0, for some $n \ge n_0$ S_n has a bounded continuous density f_n with respect to Lebesque measure on R^d, then as $n \to \infty$

$$f_n(nx) \sim (2\pi n)^{-d/2}|\Sigma(x)|^{-1/2} \cdot \exp[-n\phi(x)],$$

where $\Sigma(x) = \nabla^2\psi(\hat\theta(x))$, the covariance matrix of X_1 under $P_{\hat\theta(x)}$.

Proof: See Borovkov and Rogozin (1965).

By Proposition 1

$$f_{m-\tau}(\xi - S_\tau) \sim [2\pi(m-\tau)]^{-d/2}\left|\Sigma\left(\frac{\xi - S_\tau}{m-\tau}\right)\right|^{-1/2} \exp\left[-(m-\tau)\phi\left(\frac{\xi - S_\tau}{m-\tau}\right)\right]$$
$$f_m(\xi) \sim (2\pi m)^{-d/2}|\Sigma(\xi_0)|^{-1/2}e^{-m\phi(\xi_0)}$$

so

$$\begin{aligned}E\left\{\frac{f_{m-\tau}(\xi - S_\tau)}{f_m(\xi)}\right\} &\sim E_0\left\{\left(\frac{m}{m-\tau}\right)^{d/2}|\Sigma(\xi_0)|^{1/2}\cdot\left|\Sigma\left(\frac{\xi - S_\tau}{m-\tau}\right)\right|^{-1/2}\right.\\ &\quad\left.\cdot\exp\left[m\phi(\xi_0) - (m-\tau)\phi\left(\frac{\xi - S_\tau}{m-\tau}\right)\right]\right\}\\ &= (1-t_{\xi_0})^{-d/2}|\Sigma(\xi_0)|^{1/2}\cdot\left|\Sigma\left(\frac{\xi_0}{1-t_{\xi_0}}\right)\right|^{-1/2}\exp[-m(a_0 - \phi(\xi_0))]\cdot E_0\{e^{-R_m(\tau)}\}\end{aligned}$$

where $R_m(\tau) = (m-\tau)\phi\left(\frac{\xi - S_\tau}{m-\tau}\right) - a$ is the excess over the boundary at stopping time τ. Let $K_\xi(t,x) = (1-t)\phi\left(\frac{\xi_0 - x}{1-t}\right)$. Theorem 2 of Chapter 3 of Hogan (1984) asserts that as $m \to \infty$

Observe that

$$\begin{aligned}\nabla\phi^{(2)}(\xi_0^{(2)}) &= \nabla[\xi_0^{(2)}\cdot\theta_0^{(2)} - \psi^{(2)}(\theta_0^{(2)})] \\ &= \theta_0^{(2)} + (\xi_0^{(2)} - \nabla_{\theta^{(2)}}\psi^{(2)}(\theta_0^{(2)}))\cdot J(\theta_0^{(2)}\to\xi_0^{(2)}) \\ &= \theta_0^{(2)}\end{aligned}$$

where $J(\theta_0^{(2)}\to\xi_0^{(2)})$ is the Jacobian matrix of the mapping $\theta_0^{(2)}\to\xi_0^{(2)}$. So

$$\begin{aligned}(m-n)\phi^{(2)}[(\xi^{(2)} - S_n^{(2)})/m-n] &= (m-n)\phi^{(2)}(\xi_0^{(2)}) + (n\xi_0^{(2)} - S_n^{(2)})\theta_0^{(2)} + o(1) \quad \text{a.s.} \\ &= m\phi^{(2)}(\xi_0^{(2)}) - n[\xi_0^{(2)}\theta_0^{(2)} - \psi^{(2)}(\theta_0^{(2)})] + o(1) \quad \text{a.s..}\end{aligned}$$

Substituting the equation above into (4.1) completes the proof.

Before we investigate the nonlinear boundary crossing problem let us consider the linear problem first. Suppose we want to find the excess over the hyperplane $\mathcal{H} = \{x : \gamma\cdot x = c\}$ by the d-dimensional random walk S_n in the norm direction γ, where γ is a d-dimensional vector satisfying $E(\gamma\cdot S_1) = \gamma\cdot\mu > 0$.

The problem above can be converted into a one-dimensional problem. In fact the first time S_n crosses $\mathcal{H}$ is the same as the first time the one-dimensional random walk $\gamma\cdot S_n$ crosses the constant level c. Moreover the excess in the normal direction by S_n is exactly $(\gamma\cdot S_{T_c} - c)$ where $T_c = \inf\{n,\ \gamma\cdot S_n > c\}$.

The next theorem relates the excess over the linear boundary of a conditional random walk to that of an independent random walk.

Theorem 2.

$$\lim_{m\to\infty} P_{\xi^{(2)}}^{(m)}\{\gamma\cdot S_{T_c} - c \le x\} = P_{(0,\theta_0^{(2)})}\{\gamma\cdot S_{T_c} - c \le x\}.$$

Moreover the equality above still holds when c is allowed to depend on m in such a way that $0(m^{1/2-\alpha}) = c \to \infty$ where α is any positive number less than 1/2.

Proof: Clearly $P_{\xi^{(2)}}^{(m)}\{T_c \le m^{(1-\alpha)/2}\} \to 1$ as $m\to\infty$. So for any given $\epsilon > 0$ we can find m_1 such that $\forall m > m_1$ the inequality below holds:

$$|P_{\xi^{(2)}}^{(m)}\{\gamma\cdot S_{T_c} - c \le x;\ T_c \le m^{(1-\alpha)/2}\} - P_{\xi^{(2)}}^{(m)}\{\gamma\cdot S_{T_c} - c \le x\}| < \epsilon.$$

Chapter 4

Nonlinear Renewal Theory for Conditional Random Walks

Here we study the asymptotic distribution of excess over the boundary by a conditional random walk. The notation used here is consistent with that of Chapters 1, 2, and 3, except where stated otherwise. We begin with the following lemma which shows that as the time of conditioning becomes remote a conditional random walk behaves like an independent one locally. Let $\xi_0^{(2)} = \xi^{(2)} \cdot m^{-1}$, $\phi^{(2)}(x) = \sup_{\theta^{(1)}=0}[\theta' x - \psi(\theta)]$, $\psi^{(2)}(\theta) = \psi[(0, \theta^{(2)})]$. For the sake of simplicity write $\theta_0^{(2)}$ for $\theta^{(2)}(\xi_0^{(2)})$.

Lemma 1. If $n = o(m^{1/2})$ then

$$\lim_{m\to\infty} \frac{dP^{(m)}_{\xi^{(2)}}}{dP_{(0,\theta_0^{(2)})}}(S_1, \cdots, S_n) = 1 \quad \text{a.s. } P_{(0,\theta_0^{(2)})}.$$

Proof:

$$\begin{aligned}\frac{dP^{(m)}_{\xi^{(2)}}}{dP_{(0,\theta_0^{(2)})}}(S_1, \cdots, S_n) &= [f(S_1) \cdot f(S_2 - S_1) \cdots f(S_n - S_{n-1}) f^{(2)}_{m-n}(\xi^{(2)} - S_n^{(2)}) / f^{(2)}_m(\xi^{(2)})] \\ &\cdot [f_{(0,\theta_0^{(2)})}(S_1) \cdot f_{(0,\theta_0^{(2)})}(S_2 - S_1) \cdots f_{(0,\theta_0^{(2)})}(S_n - S_{n-1})]^{-1} \\ &= f^{(2)}_{m-n}(\xi^{(2)} - S_n^{(2)}) \cdot [f^{(2)}_m(\xi^{(2)}) \cdot \exp\{S_n^{(2)} \cdot \theta_0^{(2)} - n\psi^{(2)}(\theta_0^{(2)})\}]^{-1}.\end{aligned}$$

By Proposition 1 of Chapter 2 the quantity above tends to

$$\begin{aligned}&[m/(m-n)]^{d_2/2} |\Sigma^{(2)}(\xi_0^{(2)})|^{1/2} |\Sigma^{(2)}[(\xi^{(2)} - S_n^{(2)})/(m-n)]|^{-1/2} \\ &\cdot \exp\{-(m-n)\phi^{(2)}[(\xi^{(2)} - S_n^{(2)})/(m-n)] \\ &\cdot \exp[m\phi^{(2)}(\xi_0^{(2)})] \cdot \exp\{-S_n^{(2)} \cdot \theta_0^{(2)} + n\psi^{(2)}(\theta_0^{(2)})\}.\end{aligned} \tag{4.1}$$

Now

$$\phi^{(2)}[(\xi^{(2)} - S_n^{(2)})/(m-n)] = \phi^{(2)}(\xi_0^{(2)}) + [(n\xi_0^{(2)} - S_n^{(2)})/(m-n)] \cdot \nabla\phi^{(2)}(\xi_0^{(2)}) + o(m^{-1}) \quad \text{a.s.}$$

Observe that

$$\begin{aligned}\nabla\phi^{(2)}(\xi_0^{(2)}) &= \nabla[\xi_0^{(2)}\cdot\theta_0^{(2)} - \psi^{(2)}(\theta_0^{(2)})]\\ &= \theta_0^{(2)} + (\xi_0^{(2)} - \nabla_{\theta^{(2)}}\psi^{(2)}(\theta_0^{(2)}))\cdot J(\theta_0^{(2)}\to\xi_0^{(2)})\\ &= \theta_0^{(2)}\end{aligned}$$

where $J(\theta_0^{(2)}\to\xi_0^{(2)})$ is the Jacobian matrix of the mapping $\theta_0^{(2)}\to\xi_0^{(2)}$. So

$$\begin{aligned}(m-n)\phi^{(2)}[(\xi^{(2)} - S_n^{(2)})/m-n] &= (m-n)\phi^{(2)}(\xi_0^{(2)}) + (n\xi_0^{(2)} - S_n^{(2)})\theta_0^{(2)} + o(1) \quad \text{a.s.}\\ &= m\phi^{(2)}(\xi_0^{(2)}) - n[\xi_0^{(2)}\theta_0^{(2)} - \psi^{(2)}(\theta_0^{(2)})] + o(1) \quad \text{a.s.}.\end{aligned}$$

Substituting the equation above into (4.1) completes the proof.

Before we investigate the nonlinear boundary crossing problem let us consider the linear problem first. Suppose we want to find the excess over the hyperplane $\mathcal{H} = \{x : \gamma\cdot x = c\}$ by the d-dimensional random walk S_n in the norm direction γ, where γ is a d-dimensional vector satisfying $E(\gamma\cdot S_1) = \gamma\cdot\mu > 0$.

The problem above can be converted into a one-dimensional problem. In fact the first time S_n crosses $\mathcal{H}$ is the same as the first time the one-dimensional random walk $\gamma\cdot S_n$ crosses the constant level c. Moreover the excess in the normal direction by S_n is exactly $(\gamma\cdot S_{T_c} - c)$ where $T_c = \inf\{n,\ \gamma\cdot S_n > c\}$.

The next theorem relates the excess over the linear boundary of a conditional random walk to that of an independent random walk.

Theorem 2.

$$\lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\{\gamma\cdot S_{T_c} - c \le x\} = P_{(0,\theta_0^{(2)})}\{\gamma\cdot S_{T_c} - c \le x\}.$$

Moreover the equality above still holds when c is allowed to depend on m in such a way that $0(m^{1/2-\alpha}) = c\to\infty$ where α is any positive number less than 1/2.

Proof: Clearly $P^{(m)}_{\xi^{(2)}}\{T_c \le m^{(1-\alpha)/2}\}\to 1$ as $m\to\infty$. So for any given $\epsilon > 0$ we can find m_1 such that $\forall m > m_1$ the inequality below holds:

$$|P^{(m)}_{\xi^{(2)}}\{\gamma\cdot S_{T_c} - c\le x;\ T_c \le m^{(1-\alpha)/2}\} - P^{(m)}_{\xi^{(2)}}\{\gamma\cdot S_{T_c} - c \le x\}| < \epsilon.$$

Chapter 4

Nonlinear Renewal Theory for Conditional Random Walks

Here we study the asymptotic distribution of excess over the boundary by a conditional random walk. The notation used here is consistent with that of Chapters 1, 2, and 3, except where stated otherwise. We begin with the following lemma which shows that as the time of conditioning becomes remote a conditional random walk behaves like an independent one locally. Let $\xi_0^{(2)} = \xi^{(2)} \cdot m^{-1}$, $\phi^{(2)}(x) = \sup_{\theta^{(1)}=0}[\theta' x - \psi(\theta)]$, $\psi^{(2)}(\theta) = \psi[(0, \theta^{(2)})]$. For the sake of simplicity write $\theta_0^{(2)}$ for $\theta^{(2)}(\xi_0^{(2)})$.

Lemma 1. If $n = o(m^{1/2})$ then

$$\lim_{m\to\infty} \frac{dP^{(m)}_{\xi^{(2)}}}{dP_{(0,\theta_0^{(2)})}}(S_1, \cdots, S_n) = 1 \quad \text{a.s. } P_{(0,\theta_0^{(2)})}.$$

Proof:

$$\begin{aligned}
\frac{dP^{(m)}_{\xi^{(2)}}}{dP_{(0,\theta_0^{(2)})}}(S_1, \cdots, S_n) &= [f(S_1) \cdot f(S_2 - S_1) \cdots f(S_n - S_{n-1}) f^{(2)}_{m-n}(\xi^{(2)} - S_n^{(2)}) / f^{(2)}_m(\xi^{(2)})] \\
&\quad \cdot [f_{(0,\theta_0^{(2)})}(S_1) \cdot f_{(0,\theta_0^{(2)})}(S_2 - S_1) \cdots f_{(0,\theta_0^{(2)})}(S_n - S_{n-1})]^{-1} \\
&= f^{(2)}_{m-n}(\xi^{(2)} - S_n^{(2)}) \cdot [f^{(2)}_m(\xi^{(2)}) \cdot \exp\{S_n^{(2)} \cdot \theta_0^{(2)} - n\psi^{(2)}(\theta_0^{(2)})\}]^{-1}.
\end{aligned}$$

By Proposition 1 of Chapter 2 the quantity above tends to

$$\begin{aligned}
&[m/(m-n)]^{d_2/2} |\Sigma^{(2)}(\xi_0^{(2)})|^{1/2} |\Sigma^{(2)}[(\xi^{(2)} - S_n^{(2)})/(m-n)]|^{-1/2} \\
&\cdot \exp\{-(m-n)\phi^{(2)}[(\xi^{(2)} - S_n^{(2)})/(m-n)] \\
&\cdot \exp[m\phi^{(2)}(\xi_0^{(2)})] \cdot \exp\{-S_n^{(2)} \cdot \theta_0^{(2)} + n\psi^{(2)}(\theta_0^{(2)})\}.
\end{aligned} \tag{4.1}$$

Now

$$\phi^{(2)}[(\xi^{(2)} - S_n^{(2)})/(m-n)] = \phi^{(2)}(\xi_0^{(2)}) + [(n\xi_0^{(2)} - S_n^{(2)})/(m-n)] \cdot \nabla\phi^{(2)}(\xi_0^{(2)}) + o(m^{-1}) \quad \text{a.s.}$$

exist, we have $\tau/m \to t_{\xi_0}$ with $P^{(m)}_{\xi^{(2)}}$ probability 1, and $\lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\{\tau \le m - m_0\} = 1$.

Combining all the results above, we find that (3.4) is approximated by

$$(1 - t_{\xi_0})^{(d_1/2}|\Sigma(\xi_0)|^{1/2} \cdot \left|\Sigma\left(\frac{\xi^{(1)}}{1 - t_{\xi_0}}, \xi_0^{(2)}\right)\right|^{-1/2} E^{(m)}_{\xi^{(2)}}\{e^{-R_m}\} \exp\{-m[a_0 - \Lambda(\xi_0)]\}, \quad (3.5)$$

where $R_m = (m - \tau)\Lambda\left(\frac{\xi - S_\tau}{m-\tau}\right) - a$ is the excess over the boundary at the stopping time τ. To finish the approximation we need to identify $E^{(m)}_{\xi^{(2)}}\{e^{-R_m}\}$, the excess over the boundary by a conditional random walk. It seems to me that this topic has not been treated in the literature before. In the next chapter we shall study this topic and give some applications.

where $E^{(m)}_{\xi^{(2)}}$ is the expectation corresponding to the conditional probability $P^{(m)}_{\xi^{(2)}}$.

By Proposition 1 of Chapter 2 we have

$$\begin{aligned}
\frac{f_{m-r}(\xi - S_r)}{f^{(2)}_{m-r}(\xi^{(2)} - S^{(2)}_r)} &\sim [2\pi(m-r)]^{-d_1/2} \left|\Sigma^{(2)}\left(\frac{\xi^{(2)} - S^{(2)}_r}{m-r}\right)\right|^{1/2} \left|\Sigma\left(\frac{\xi - S_r}{m-r}\right)\right|^{-1/2} \\
&\exp\left\{-(m-r)\left[\phi\left(\frac{\xi - S_r}{m-r}\right) - \phi_0\left(\frac{\xi^{(2)} - S^{(2)}_r}{m-r}\right)\right]\right\} \\
&= [2\pi(m-r)]^{-d_1/2} \left|\Sigma^{(2)}\left(\frac{\xi^{(2)} - S^{(2)}_r}{m-r}\right)\right|^{1/2} \cdot \left|\Sigma\left(\frac{\xi - S^{(2)}_r}{m-r}\right)\right|^{-1/2} \\
&\exp\left\{-(m-r)\Lambda\left(\frac{\xi - S_r}{m-r}\right)\right\},
\end{aligned} \tag{3.2}$$

$$\begin{aligned}
\frac{f^{(2)}_m(\xi^{(2)})}{f_m(\xi)} &\sim (2\pi m)^{d_1/2} |\Sigma(\xi_0)|^{1/2} \cdot |\Sigma^{(2)}(\xi^{(2)}_0)|^{-1/2} \exp\{m[\phi(\xi_0) - \phi_0(\xi^{(2)}_0)]\} \\
&= (2\pi m)^{d_1/2} |\Sigma(\xi_0)|^{1/2} \cdot |\Sigma^{(2)}(\xi^{(2)}_0)|^{-1/2} e^{m\Lambda(\xi_0)}
\end{aligned} \tag{3.3}$$

where $\Sigma^{(2)}(\mu)$ is the covariance matrix of $S^{(2)}_1$ under P_μ.

Substituting (3.2), (3.3) into (3.1) we have

$$\begin{aligned}
E^{(m)}_{\xi^{(2)}}\Bigg\{ \left(\frac{m}{m-r}\right)^{d_1/2} &\left|\Sigma^{(2)}\left(\frac{\xi^{(2)} - S^{(2)}_r}{m-r}\right)\right|^{1/2} |\Sigma(\xi_0)|^{1/2} \cdot |\Sigma^{(2)}(\xi^{(2)}_0)|^{-1/2} \left|\Sigma\left(\frac{\xi - S_r}{m-r}\right)\right|^{-1/2} \\
&\exp\left[a - (m-r)\Lambda\left(\frac{\xi - S_r}{m-r}\right)\right];\ r \le m - m_0 \Bigg\} \cdot \exp[-m(a_0 - \Lambda(\xi_0))].
\end{aligned} \tag{3.4}$$

It is not hard to show that with high probability $S^{(2)}_r \sim r \cdot \xi^{(2)}_0$ under $P^{(m)}_{\xi^{(2)}}$, so

$$\Sigma^{(2)}\left(\frac{\xi^{(2)} - S^{(2)}_r}{m-r}\right) \to \Sigma^{(2)}\left(\frac{\xi^{(2)} - r \cdot \xi^{(2)}_0}{m-r}\right) = \Sigma^{(2)}(\xi^{(2)}_0).$$

Similarly,

$$S_r \sim r(0, \xi^{(2)}_0) \cdot \Sigma\left(\frac{\xi - S_r}{m-r}\right) \to \Sigma\left(\frac{\xi^{(1)}_0}{1 - t_{\xi_0}}, \xi^{(2)}_0\right)$$

assumes $r/m \to t_{\xi_0}$ under $P^{(m)}_{\xi^{(2)}}$.

For those ξ_0 such that

$$t_{\xi_0} = \inf\left\{t : 0 < t < 1 - m_0/m,\ (1-t)\Lambda\left(\frac{\xi^{(1)}_0}{1-t},\ \xi^{(2)}_0\right) = a_0\right\}$$

Define

$$Q(A) = \int P^{(m)}_{\lambda^0,\xi}(A)\frac{f_m(\xi-\lambda^0)}{f^{(2)}_m(\xi^{(2)})}d\lambda_{(1)}$$

where $\lambda^0 = (\lambda^{(1)}, \mathbf{o})$ where $\lambda^{(1)} \in R^{d_1}$, $\mathbf{o}$ is the zero vector in R^{d_2}, $d_1 + d_2 = d$, $f^{(2)}_m$ is the density for $S^{(2)}_n$ under P_0 where

$$S_n = (S^{(1)}_n, S^{(2)}_n), \quad S^{(1)}_n \in R^{d_1}, \quad S^{(2)}_n \in R^{d_2}.$$

Now the likelihood ratio

$$\frac{dQ}{dP^{(m)}_\xi}(S_n,\cdots,S_m) = \int \frac{dP^{(m)}_{\lambda^0,\xi}}{dP^{(m)}_{0,\xi}}(S_n,\cdots,S_m)\frac{f_m(\xi-\lambda^0)}{f^{(2)}_m(\xi^{(2)})}d\lambda^{(1)} = \frac{f_m(\xi)f^{(2)}_n(S^{(2)}_n)}{f^{(2)}_n(\xi^{(2)})f_n(S_n)}.$$

Observe that

$$\frac{f_m(\xi-\lambda^0)}{f^{(2)}_m(\xi^{(2)})} = f_m(\xi-\lambda^0 \mid \xi^{(2)})$$

is the conditional density of S_m given $S^{(2)}_m$.

The Q distribution of S_n, $n = m, m-1, \cdots$ running backward from $S_m = m\xi_0$ is the same as the conditional P_0 distribution of $m\xi_0 - S_n$, $n = 0, 1, \cdots$ running forward and tied down at $S^{(m)}_m = m\xi^{(2)}_0$. Under the reverse time scale T^* is a stopping time so the Wald likelihood ratio identity gives

$$\begin{aligned} P^{(m)}_\xi(T < m) = P^{(m)}_\xi(T^* \geq m_0) &= E_Q\left\{\frac{dP^{(m)}_\xi}{dQ}(S_{T^*},\cdots,S_m);\ T^* \geq m_0\right\} \\ &= E_Q\left\{\frac{f_{T^*}(S_{T^*})}{f^{(2)}_{T^*}(S^{(2)}_{T^*})}\frac{f^{(2)}_m(\xi^{(2)})}{f_m(\xi)};\ T^* \geq m_0\right\} \end{aligned}$$

Let

$$\tau = \inf\left\{n : (m-n)\Lambda\left(\frac{\xi - S_n}{m-n}\right) > a,\ n \leq m - m_0\right\}.$$

It is easy to see that the distribution of T^* under Q is the same as the distribution of $m-\tau$ under $P^{(m)}_{\xi^{(2)}}(A) = P_0(A \mid S^{(2)}_m = \xi^{(2)})$. So the expectation above can be replaced by

$$E^{(m)}_{\xi^{(2)}}\left\{\frac{f_{m-\tau}(\xi - S_\tau)}{f^{(2)}_{m-\tau}(\xi^{(2)} - S^{(2)})_\tau)}\frac{f^{(2)}_m(\xi^{(2)})}{f_m(\xi)};\ \tau \leq m - m_0\right\} \tag{3.1}$$

Chapter 3

The Composite Null Hypothesis Case

In this chapter we consider the more difficult problem of composite hypotheses. The null hypotheses considered in this chapter are of the form

$$H_0 : \theta \in \Theta_0 = \{\theta : \theta^{(1)} = 0\},$$

where $\theta = (\theta^{(1)}, \theta^{(2)})$ $\theta^{(1)} = (\theta_1, \cdots, \theta_{d_1}) \in R^{d_1}$, $\theta^{(2)} = (\theta_{d_1+1}, \cdots, \theta_d)$, $d_1 + d_2 = d$. So $\theta^{(2)}$ plays the role of nuisance parameter here. The stopping rule corresponding to this case is given by

$$T = \inf\{n \geq m_0;\ n\Lambda(S_n/n) > a\} \quad \text{where} \quad \Lambda(x) = \phi(x) - \phi_0(x).$$

In the case of the composite null hypothesis the most tricky part of both the forward and backward method is to find a measure Q such that its likeihood ratio with respect to the probability measure under the null hypothesis is a simple function of the stopping rule (asymptotically). For the forward method, the Q measure is taken to be the mixture of P_θ over a submanifold N of the parameter space Θ e.g. $Q = \int_N P_\theta d\sigma_N(\theta) / \int d\sigma_N(\theta)$, where $d\sigma_N(\theta)$ is the differential element of the manifold N, see Lalley (1983) for details. For the backward method, which we will develop here, Q is defined by randomizing the starting point of the sufficient process S_n according to a conditional distribution. To be more precise, define $T^* = \sup\{n,\ n[\phi(S_n/n) - \phi_0(S_n/n)] > a\}$ and $P_{\lambda,\xi}^{(m)}(A) = P(A \mid S_0 = \lambda,\ S_m = \xi)$, $P_\xi^{(m)}(A) = P(A \mid S_m = \xi)$. It is easy to see that

$$\frac{dP_{\lambda,\xi}^{(m)}}{dP_{0,\xi}^{(m)}}(S_n, \cdots, S_m) = \frac{f_n(S_n - \lambda) f_m(\xi)}{f_n(S_n) f_m(\xi - \lambda)}.$$

Remark 1. The change of variable (2.12) amounts to saying the backward and forward process hit the boundary at the same time. This interpretation helps us find the appropriate change of variable in more complicated situations. See Chapter 5 for an example of the composite null hypothesis case.

Remark 2. The relation that the increment of forward and backward processes have the simple likelihood ratio e^y is not accidental. See Chapter 5 for another example.

Remark 3. Theorem 2 has the following merits as far as numerical computation is concerned. It relates the excess over the boundary of the forward process and the backward process in such an elegant way that no matter what method you use you only need one program to compute the excess over the boundary.

Remark 4. Theorem 2 also implies that if we ignore the excess over the boundary, in general the forward and backward methods will not give the same approximation.

Proof: By (2.15)

$$\lim_{a \to \infty} E\{\exp[-(U_{\tau_a} - a)]\} = (EU_{\tau_+})^{-1}[1 - E(e^{-U_{\tau_+}})], \tag{2.16}$$

where τ_+ is the first time the corresponding random walk is positive. When there is possibility of confusion, we also write τ_+^U to indicate which random walk we refer to. By Wald's lemma

$$E(U_{\tau_+}) = \mu_Y \cdot E\tau_+^U. \tag{2.17}$$

By the duality lemma

$$E\tau_+^U = P\{\tau_-^U = \infty\}^{-1}. \tag{2.18}$$

where τ_-^U is the first time the corresponding random walk is nonnegative. Wald's likelihood ratio identity gives

$$P(\tau_+^{-V} < \infty) = E(e^{-U_{\tau_+}}). \tag{2.19}$$

Substituting (2.17)-(2.19) into (2.16) we have

$$\lim_{a \to \infty} E\{\exp[-(U_{\tau_a} - a)]\} = P\{\tau_+^U = \infty\}P\{\tau_+^{-V} = \infty\}/\mu_Y. \tag{2.20}$$

Arguing exactly the same way we obtain

$$\lim_{a \to \infty} E\{\exp[-(V_{\tau_a} - a)]\} = P\{\tau_+^{-U} = \infty\}P\{\tau_-^V = \infty\}/\mu_Z. \tag{2.21}$$

By obvious scale property

$$P\{\tau_+^{-U} = \infty\} = P\{\tau_-^U = \infty\} \tag{2.22}$$

$$P\{\tau_+^{-V} = \infty\} = P\{\tau_-^V = \infty\}. \tag{2.23}$$

Dividing (2.20) by (2.21), using (2.22) and (2.23), we get the desired result. This completes the proof of Theorem 2.

It is routine to check that

$$\frac{\mu_Z}{\mu_Y} = \frac{\psi(\theta)}{\theta\psi'(\theta) - \psi(\theta)} = 1 - \frac{\theta\psi'(\theta)}{\phi(\psi'(\theta))}.$$

This proves (2.14). Several remarks are called for.

By a nonlinear renewal theorem of Lai and Siegmund (1977)

$$\lim_{m\to\infty} E_\theta\{\exp[-T\phi(S_T/T)-a)\} = [E(U_{\tau_+})]^{-1}\int_0^\infty e^{-x}P\{U_{\tau_+}>x\}dx,$$

where $Y = \phi(\psi'(\theta)) + (X_1 - \psi'(\theta))\theta = \theta X_1 - \psi(\theta)$, X_1 is distributed according to P_θ and $U_n = \sum_{i=1}^n Y_i$ Y_i i.i.d. and each Y_i has the same distribution as Y. Integrating by parts gives

$$(E\,U_{\tau_+})^{-1}\int_0^\infty e^{-x}P\{U_{\tau_+}>x\}dx = (EU_{\tau_+})^{-1}[1-E(e^{-U_{\tau_+}})]. \tag{2.15}$$

The increment Z of the random walk V_n in Theorem 1 has the same distribution as

$$\begin{aligned}\nabla K_{\xi_0}(t_{\xi_0},0)\cdot(1,X_1) &= \left(\frac{\partial}{\partial t}(1-t)\phi\left(\frac{\xi_0}{1-t}\right)\Big|_{(t_{\xi_0},0)}, \frac{\partial}{\partial x}(1-t)\phi\left(\frac{\xi_0-x}{1-t}\right)\Big|_{(t_{\xi_0},0)}\right)\\ \cdot(1,X_1) &= -\left[\phi\left(\frac{\xi_0}{1-t_{\xi_0}}\right)+\frac{\xi_0}{1-t_{\xi_0}}\phi\left(\frac{\xi_0}{1-t_{\xi_0}}\right)-X_1\phi\left(\frac{\xi_0}{1-t-\xi_0}\right)\right]\\ &= -[\phi(\psi'(\theta))+\psi'(\theta)\phi'(\psi'(\theta))-X_1\phi'(\psi'(\theta))] \quad \text{by (2.11)}\\ &= -(\theta X_1-\psi(\theta))\end{aligned}$$

where we have used the identities $\phi(\psi'(\theta)) = \theta\psi'(\theta) - \psi(\theta)$ and $\phi'(\psi'(\theta)) = \theta$ in the third equation above.

The likelihood ratio of Y with respect to $-Z$ is equal to

$$\frac{f_Y(y)}{F_{-Z}(y)} = \frac{\exp\left[\theta\left(\frac{y+\psi(\theta)}{\theta}\right)-\psi(\theta)\right]f_0\left(\frac{y+\psi(\theta)}{\theta}\right)}{f_0\left(\frac{y+\psi(\theta)}{\theta}\right)} = e^y.$$

The following theorem is all we need to complete the program.

Theorem 2. Let $Y_1, Y_2, \cdots$ and $Z_1, Z_2, \cdots$ be two sequences of independent identically distributed random variables. Also let $U_n = \sum_{i=1}^n Y_i$, $V_n = \sum_{i=1}^n Z_i$. If the likelihood ratio of Y_i with respect to $-Z_i$ is equal to e^y then

$$\frac{\lim_{a\to\infty} E\{\exp[-U_{\tau_a}-a)]\}}{\lim_{a\to\infty} E\{\exp[-(V_{\tau_a}-a)]\}} = \frac{\mu_Z}{\mu_Y}$$

where τ_a is the first time the corresponding random walk exceeds a, $\mu_X = EX_1 > 0$, $\mu_Y = EY_1 > 0$.

Expanding the likelihood ratio $\frac{dQ}{dP_0}$ about $\hat{\theta}$ and using Laplace's method we have

$$\begin{aligned}\frac{dQ}{dP_0}(X_1,\cdots,X_n) \equiv L_n &= \int \exp[\ell_n(\theta)-\ell_n(0)](2\pi)^{-1/2}d\theta \\ &\sim \exp[\ell_n(\hat{\theta})-\ell_n(0)]\exp[\ddot{\ell}_n(\hat{\theta})(\theta-\hat{\theta}_n)^2/2](2\pi)^{-1/2}d\theta \qquad (2.8)\\ &\sim \exp[\phi(S_n/n)]/|\ddot{\ell}_n(\hat{\theta}_n)|^{1/2}.\end{aligned}$$

Substitute (2.3), (2.4), (2.6)–(2.8) into (2.2). We have

$$P_0\{m_0 < t \le m,\ \phi(S_m/m) \le c/m\}$$
$$\sim \int_{\{J(\theta,0)<\eta_0,\phi(\bar{t}_\theta\psi'(\theta))<c/m\}} \left[\frac{I(\theta)}{J(\theta,0)}\right]^{1/2} E_\theta[e^{-(\phi(S_T/T)-a)}]d\theta(2\pi)^{-1/2}\cdot a^{1/2}e^{-a} \qquad (2.9)$$

By Corollary 1 the backward method gives us

$$P_0\{m_0 \le T < m,\ \phi(S_m/m) \le c/m\} \qquad (2.10)$$
$$\sim (m/2\pi)^{1/2}e^{-a}\int_{\{\phi(\xi_0)<c/m,\ \phi(\xi_0/1-t_{\xi_0})\le\eta_0\}} \left|\ddot{\Phi}\left(\frac{\xi_0}{1-t_{\xi_0}}\right)\right|^{-1/2}(1-t_{\xi_0})^{-1/2}\nu_-(\xi_0)d\xi_0.$$

To show (2.9) agrees with (2.10), let us make the change of variable

$$\frac{\xi_0}{1-t_{\xi_0}} = \psi'(\theta). \qquad (2.11)$$

By (2.1) and (2.5) we have

$$\bar{t}_\theta = 1 - t_{\xi_0} \qquad (2.12)$$

Now

$$\frac{d}{d\theta}\bar{t}_\theta = \frac{d}{d\theta}\frac{\eta_1}{J(\theta,0)} = \frac{-\eta_1(\theta)\psi''(\theta)}{(\theta\psi'(\theta)-\psi(\theta))^2}.$$
$$\frac{d\xi_0}{d\theta} = \frac{d}{d\theta}\bar{t}_\theta\psi'(\theta) = \psi''(\theta)\bar{t}_\theta + \psi'(\theta)\frac{d}{d\theta}\bar{t}_\theta = \psi''(\theta)\bar{t}_\theta\left[1-\frac{\theta\psi'(\theta)}{\phi(\psi'(\theta))}\right]. \qquad (2.13)$$

Substitute (2.11), (2.12), (2.13) into (2.10). We have

$$P_0\{m_0 < T \le m,\ \phi(S_m/m) < c/m\}$$
$$\sim (a/2\pi)^{1/2}e^{-a}\int_{\phi(\bar{t}_\theta\psi'(\theta))<c/m,\ J(\theta,0)<\eta_0\}} \left(\frac{I(\theta)}{J(\theta,0)}\right)^{1/2}\nu_-(\bar{t}_\theta\psi'(\theta))\left|1-\frac{\theta\psi'(\theta)}{\phi(\psi'(\theta))}\right|d\theta.$$

To show (2.9) agrees with (2.10) it is sufficient to prove

$$\nu_-(\bar{t}_\theta\psi'(\theta))\left|1-\frac{\theta\psi'(\theta)}{\phi(\psi'(\theta))}\right| = \lim_{m\to\infty}E_\theta\{\exp[-(T\phi(S_T/T)-a)]\}. \qquad (2.14)$$

$E_{\theta_1}\{\log(f_{\theta_1}/f_{\theta_2})\} = (\theta_1 - \theta_2)\mu(\theta_1) - (\psi(\theta_1) - \psi(\theta_2))$, also let

$$I(\theta) = E_\theta\left\{-\frac{d^2}{d\theta^2}\log f_\theta\right\} = \psi''(\theta) = \mathrm{Var}_\theta(X_1)$$

be the Fisher information. Now define $Q(A) = \int_{-\infty}^{\infty} P_\theta(A)(2\pi)^{-1/2}d\theta$. Then

$$P_0\{m_0 < t \le m,\ \phi(S_m/m) < c/m\} = E_0\{P_0[\phi(S_m/m) \le c/m \mid \mathcal{E}_T];\ m_0 < T \le m\}$$

where $\mathcal{E}_T$ is the σ-field generated by all events A satisfying $A \cap [T = n] \in \mathcal{F}(X_1, \cdots, X_n)$ (see Chung ( )). Applying Wald's likelihood ratio identity, the expectation above is equal to

$$\begin{aligned} &E_Q\left\{\frac{dP_0}{dQ}(S_1,\cdots,S_T)P_0\{\phi(S_m/m) \le c/m \mid \mathcal{E}_T\},\ m_0 < T \le m\right\} \\ &\quad = \int E_\theta\left\{\frac{dP_0}{dQ}(S_1,\cdots,S_T)P_0[\phi(S_m/m) \le c/m \mid \mathcal{E}_T];\ m_0 < T \le m\right\}d\theta. \end{aligned} \tag{2.2}$$

It is easy to see that $P_\theta\{\lim_{n\to\infty}\phi(S_n/n) = J(\theta,0)\} = 1$, so

$$P_\theta\{\lim_{a\to\infty} a^{-1}T = [J(\theta,0)]^{-1}]\} = 1. \tag{2.3}$$

If m_0, m and a are related by $m^{-1}a = \eta_1 < m_0^{-1}a = \eta_0$ then it follows that

$$P_\theta\{m_0 < T \le m\} \to \begin{cases} 1 & \text{if } \eta_1 < J(\theta,0) < \eta_0 \\ 0 & \text{if } J(\theta,0) \notin [\eta_1, \eta_0]. \end{cases} \tag{2.4}$$

In this case we have

$$P_\theta\{\lim_{m\to\infty} m^{-1}T = \bar{t}_\theta\} = 1$$

where

$$\bar{t}_\theta = \eta_1/J(\theta,0). \tag{2.5}$$

If we impose further regularity conditions such that $\lim_{n\to\infty} n^{-1}\ddot{\ell}_n(\hat{\theta}_n) \to I(\theta)$ holds then

$$P_\theta\{\lim_{a\to\infty} a^{-1}[-\ddot{\ell}_T(\hat{\theta}_T)] = I(\theta)/J(\theta,0)\} = 1. \tag{2.6}$$

Taylor's expansion gives $\phi(S_m/m) = \phi(T\bar{X}_T/m) + \frac{1}{m}(S_m - S_T)\phi'(T\bar{X}_T/m) + o_p(1)$ under P_0, so as $m \to \infty$,

$$P_0\{\phi(S_m/m) < c/m \mid \mathcal{E}_T\} \to \begin{cases} 1 & \text{if } \phi(\bar{t}_\theta\psi'(\theta)) < c/m \\ 0 & \text{if } \phi(\bar{t}_\theta\psi'(\theta)) \ge c/m. \end{cases} \tag{2.7}$$

R_m tends to the same distribution as the excess over the boundary by the random walk $\nabla K_\xi(t_{\xi_0}, 0) \cdot \tilde{S}_n$ over a distant constant boundary, where $\tilde{S}_n = (1, S_n)$. Now the ordinary renewal theorem can be used to identify this asymptotic distribution. This completes the proof of

Theorem 1. Let $\xi = \xi_0 m$, $m\phi(\xi_0) < a$. If a, m tend to ∞ in such a way that $am^{-1} \to a_0$ and if t_{ξ_0} defined by (2.1) satisfies $0 < t_{\xi_0} < 1 - m_0/m$ then the following asymptotic relation holds

$$P_\xi^{(m)}(T \le m) \sim (1 - t_{\xi_0})^{-d/2} |\ddagger(\xi_0)|^{1/2} \cdot \left| \ddagger\left(\frac{\xi_0}{1 - t_{\xi_0}}\right)\right|^{-1/2} \exp[-m(a_0 - \phi(\xi_0))]\nu_-(\xi_0)$$

where $\nu_-(\xi_0) = [E(V_{\tau_+})]^{-1} \int_0^\infty e^{-x} P\{V_{\tau_+} > x\} dx$, $V_n = \sum_{i=1}^n Z_i$ Z_i are i.i.d. and have the same distribution as $\nabla K_{\xi_0}(t_{\xi_0}, 0) \cdot (1, X_1)$, X_1 is distributed according to F_0, and $\tau_+ = \inf\{n > 0,\ V_n > 0\}$

Corollary 1.

$$\begin{aligned} &P_0\{T < m,\ m\phi(S_m/m) \le c\} \\ &\quad \sim (m/2\pi)^{d/2} \int_{\{\phi(\xi_0) \le c_0, t_{\xi_0} < 1 - m_0/m\}} (1 - t_{\xi_0})^{-1/2} \nu_-(\xi_0) d\xi_0 \cdot e^{-c} \end{aligned}$$

where $c_0 = cm^{-1}$.

Corollary 2.

$$\begin{aligned} &P_\theta\{T < m,\ m\phi(S_m/m) \le c\} \\ &\quad \sim (m/2\pi)^{d/2} \exp\{-m[a_0 - \psi(\theta)]\} \int_{\{\phi(\xi_0) \le c_0, 0 < t_0 \le 1 - m_0/m\}} (1 - t_{\xi_0})^{-1/2} \nu_-(\xi_0) e^{m\theta' \xi_0} d\xi_0 \end{aligned}$$

The proofs of Corollary 1 and 2 are omitted here. In principle it follows from integrating the result of Theorem 1. Corollary 2 needs further simplification. The actual procedure may depend on the probability model under consideration, but the arguments in Chapter 5 may provide a clue. Next we'll derive Corollary 1 in the one parameter exponential family using the forward method, then show that the forward and backward methods give the same result. Let $J(\theta_1, \theta_2)$ be the Kullback-Leibler distance between f_{θ_1} and f_{θ_2} i.e. $J(\theta_1, \theta_2) =$

By Wald's likelihood ratio identity

$$P^{(m)}_{\xi^{(2)}}\{\gamma\cdot S_{T_c}-c\le x;\ T_c\le m^{(1-\alpha)/2}\}$$
$$=E_{(0,\theta_0^{(2)})}\left\{\frac{dP^{(m)}_{\xi^{(2)}}}{dP_{(0,\theta_0^{(2)})}}(S_1,\cdots,S_{m^{(1-\alpha)/2}});\ \gamma\cdot S_{T_c}-c\le x,\ T_c\le m^{(1-\alpha)/2}\right\}.$$

Applying Lemma 1 and Scheffé's theorem (see e.g. page 184, Billingsley (1979)), the expectation above can be made arbitrarily close to $P_{(0,\theta_0^{(2)})}\{\gamma\cdot S_{T_c}-c\le x,\ T_c\le m^{(1-\alpha)/2}\}$ which in turn can be approximated by $P_{(0,\theta_0^{(2)})}\{\gamma\cdot S_{T_c}-c\le x\}$ when m is sufficiently large. That is

$$
\begin{aligned}
&|P^{(m)}_{\xi^{(2)}}\{\gamma\cdot S_{T_c}-c\le x\}-P_{(0,\theta_0^{(2)})}\{\gamma\cdot S_{T_c}-c\le x\}|\\
&\quad\le|P^{(m)}_{\xi^{(2)}}\{\gamma\cdot S_{T_c}-c\le x\}-P^{(m)}_{\xi^{(2)}}\{\gamma\cdot S_{T_c}-c\le x,\ T_c\le m^{(1-\alpha)/2}\}|\\
&\quad+|P^{(m)}_{\xi^{(2)}}\{\gamma\cdot S_{T_c}-c\le x,\ T_c\le m^{(1-\alpha)/2}\}-P_{(0,\theta_0^{(2)})}\{\gamma\cdot S_{T_c}-c\le x,\ T_c\le m^{(1-\alpha)/2}\}\\
&\quad+|P_{(0,\theta_0^{(2)})}\{\gamma\cdot S_{T_c}-c\le x,\ T_c\le m^{(1-\alpha)/2}\}-P_{(0,\theta_0^{(2)})}\{\gamma\cdot S_{T_c}-c\le x\}|\\
&\quad\le 3\epsilon
\end{aligned}
$$

for m sufficiently large. This completes the proof.

Although the theorem above gives the asymptotic distribution of excess over a local linear boundary, the problems of interest require that $c=0(m)$. For this case a "restarting argument" is needed. The restarting argument can also be used to determine the asymptotic joint distribution of stopping time and excess over the boundary. Before we get into this let us find the asymptotic marginal distribution of the stopping time first. In Lemma 4 and Theorem 5 below the setting are the same as Theorem 2 except that $c=c_0m$ for some $c_0>0$. We will need

Proposition 3 (Borisov, 1978): Let Y_i $i=1,2,\cdots$ be a sequence of i.i.d. random variables such that $EY_i=0$ Var $Y_1=1$ and the moment generating function $E(e^{tY_1})$ exist in a neighborhood of zero. Also let $V_n=\sum_{i=1}^n Y_i$, $W_n(t)$ be the random polygonal curve with vertices at $(k/n,\ V_n/\sqrt{n})$, that is

$$W_n(t)=\frac{V_k}{\sqrt{n}}+\sqrt{n}\left(t-\frac{k}{n}\right)Y_{k+1},\quad \text{if}\quad \frac{k}{n}\le t\le\frac{k+1}{n}.$$

Then the distribution of $W_n(t)$ conditioned on $W_n(1)=0$ converges to the measure generated by the Brownian bridge $W^0(t) = W(t) - tW(1)$ where $W(t)$ denotes the standard Wiener process.

Lemma 4.

$$\lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\left\{\frac{T_c - c/\mu^{(2)}_\gamma}{\sqrt{(1-\frac{c_0}{\mu^{(2)}_\gamma})\sigma^2_{(0,\xi^{(2)}_0)}c(\mu^{(2)}_\gamma)^{-3})}} < x\right\} = \Phi(x)$$

where $\mu^{(2)}_\gamma = \gamma\cdot(0,\xi^{(2)}_0)$, $\sigma^2_\alpha = \mathrm{Var}_{\theta(\alpha)}(\gamma\cdot S_1)$ is the variance of $\gamma\cdot S_1$ under $P_{\theta(\alpha)}$. $\Phi(x)$ is the distribution function of standard normal distribution.

Proof: First observe that it is sufficient to prove

$$\lim_{m\to\infty} P^{(m)}_{\xi}\left\{\frac{T_c - c/\mu_\gamma}{\sqrt{(1-\frac{c_0}{\mu_\gamma})\sigma^2_{\xi_0}\mu_\gamma^{-3}}} < x\right\} = \Phi(x)$$

where $\mu_\gamma = \gamma\cdot\xi_0$. Because $P^{(m)}_{\xi^{(2)}}$ may be obtained from $P^{(m)}_{\xi}$ by integrating out $\xi^{(1)}$, and by Proposition 1 of Chapter 2 under P_0 $\xi^{(1)}_0 = \xi^{(1)}/m$ is degenerated at zero this is true because in Chapter .. we have assumed that $\mu(0) = 0$. Now

$$\frac{\mu_\gamma T_c - c}{\sqrt{\left(1-\frac{c_0}{\mu_\gamma}\right)\sigma^2_{\xi_0}c\mu_\gamma^{-1}}} = \frac{\mu_\gamma T_c - \gamma\cdot S_{T_c}}{\sqrt{\left(1-\frac{c_0}{\mu_\gamma}\right)\sigma^2_{\xi_0}c\mu_\gamma^{-1}}} + \frac{\gamma\cdot S_{T_c} - c}{\sqrt{\left(1-\frac{c_0}{\mu_\gamma}\right)\sigma^2_{\xi_0}c\mu_\gamma^{-1}}}$$

$\gamma\cdot S_{T_c} - c$ is the excess over the boundary at the stopping time T_c. Since

$$\{\gamma\cdot S_{T_c} - c > m^\epsilon\} \subset \{T_c \notin [(1-\delta)t,(1+\delta)t]\} \cup \{\max_{(1-\delta)t\le n\le(1+\delta)t}\gamma\cdot X_n > m^\epsilon\}$$

where $t = c/\mu_\gamma$.

$$P^{(m)}_\xi\{\gamma\cdot S_{T_c} - c > m^\epsilon\} \le P^{(m)}_\xi\{T_c \notin [(1-\delta)t,(1+\delta)t]\} + P^{(m)}_\xi\{\max_{(1-\delta)t\le n\le(1+\delta)t}\gamma\cdot X_n > m^\epsilon\}$$

By Proposition 3 $\lim_{m\to\infty} P^{(m)}_\xi\{T_c \notin [(1-\delta)t,(1+\delta)t]\} \to 0$ and by Lemma 8 below

$$P^{(m)}_\xi\{\max_{(1-\delta)t\le n\le(1+\delta)t}|\gamma\cdot X_n| > m^\epsilon\} \le 2\delta t P^{(m)}_\xi\{|\gamma\cdot X_1| > m^\epsilon\} \to 0 \quad \text{as } m\to\infty.$$

So $\gamma\cdot S_{T_c} - c = o_p(m^\epsilon)$ for $\forall\epsilon > 0$. This shows that

$$\frac{\mu_\gamma T_c - c}{\sqrt{\left(1-\frac{c_0}{\mu_\gamma}\right)\sigma^2_{\xi_0}c\mu_\gamma^{-1}}}$$

has the same asymptotic distribution as

$$\frac{\mu_\gamma T_c - \gamma \cdot S_{T_c}}{\sqrt{\left(1-\frac{c_0}{\mu_\gamma}\right)\sigma^2_{\xi_0}\mu_\gamma^{-1}}}.$$

Let $W_n = \gamma \cdot S_n - (\gamma \cdot \xi_0)n/m$. Then

$$P_\xi^{(m)}\left\{\frac{\mu_\gamma T_c - \gamma \cdot S_{T_c}}{\sqrt{\left(1-\frac{c_0}{\mu_\gamma}\right)\sigma^2_{\xi_0}c\mu_\gamma^{-1}}} < x\right\} = P\left\{\frac{W_{T_c}}{\sqrt{\left(1-\frac{c_0}{\mu_\gamma}\right)\sigma^2_{\xi_0}c\mu_\gamma^{-1}}} < x \mid W_m = 0\right\}$$

Notice that

$$\{|W_{T_c} - W_t| > \epsilon m^{1/2}\} \subset \{T_c \notin [(1-\delta)t,(1+\delta)t]\} \cup \{\max_{(1-\delta)t\leq n\leq(1+\delta)T}|W_n - W_t| > \epsilon m^{1/2}\}$$

and

$$\{\max_{(1-\delta)t\leq n\leq(1+\delta)t}|W_n - W_t| > \epsilon m^{1/2}\} \subset \{\max_{(1-\delta)t\leq n\leq(1+\delta)t}|W_n - W_{(1-\delta)t}| > \frac{\epsilon}{2}m^{1/2}\}.$$

So

$$\begin{aligned} P\{|W_{T_c} - W_t| > \epsilon m^{1/2} \mid W_m = 0\} &\leq P\{T_c \notin [(1-\delta)t,(1_\delta)t] \mid W_m = 0\} \\ &+ P\{\max_{(1-\delta)t\leq n\leq(1+\delta)t}|W_n - W_{(1-\delta)t}| > \frac{\epsilon}{2}m^{1/2} \mid W_m = 0\} \end{aligned}$$

We have shown $P\{T_C \notin [(1-\delta)t,(1+\delta)t] \mid W_m = 0\} \to 0$ as $m \to \infty$. By Proposition 3

$$P\{\max_{(1-\delta)t\leq n\leq(1+\delta)t}|W_n - W_{(1-\delta)t}| > \frac{\epsilon}{2}m^{1/2} \mid W_m = 0\} \to P\{\sup_{0\leq s\leq 2\delta c_0\mu_\gamma^{-1}}|W^\circ(s)| > \frac{\epsilon}{2}\}$$

which tends to zero as $\delta \to 0$.

This shows that

$$\frac{W_{T_c}}{\sqrt{m}} = \frac{W_{T_c} - W_t}{\sqrt{m}} + \frac{W_t}{\sqrt{m}}$$

has the same asymptotic distribution as $\frac{W_t}{\sqrt{m}}$, but by Proposition 3 $\frac{W_t}{\sqrt{m}}$ has asymptotic distribution $N\left(0,\sigma^2_\xi\left(1-\frac{c_0}{\mu_\gamma}\right)\frac{c_0}{\mu_\gamma}\right)$. This proves

$$\lim_{m\to\infty} P_\xi^{(m)}\left\{\frac{T_c - c/\mu_\gamma}{\sqrt{\left(1-\frac{c_0}{\mu_\gamma}\right)\sigma^2_\xi c\mu_\gamma^{-3}}} < x\right\} = \Phi(x).$$

This completes the proof.

Theorem 5.

$$\lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\left\{\frac{T - c/\mu^{(2)}_{\gamma}}{\sqrt{\left(1-\frac{\epsilon_0}{\mu^{(2)}_{\gamma}}\right)\sigma^2_{(0,\xi^{(2)}_0)}c(\mu^{(2)}_{\gamma})^{-3}}} < x, \gamma\cdot S_{T_c} - c < y\right\}$$
$$= \Phi(x)\lim_{c\to\infty} P_{(0,\theta^{(2)}_0)}\{\gamma\cdot S_{T_c} - c < y\}$$

Proof: Define $T' = \inf\{n, \gamma\cdot S_n > c - m^{1/3}\}$. Then

$$\{T - T' > \epsilon m^{2/5}\} \subset \{\max_{0\le n\le T'+\epsilon m^{2/5}} \gamma\cdot S_n \le c\} \subset \{\max_{T'\le n\le T'+\epsilon m^{2/5}} \gamma\cdot S_n - \gamma S_{T'} \le m^{1/3}\}.$$

So

$$P^{(m)}_{\xi^{(2)}}\{T - T' > \epsilon m^{2/5}\} \le P^{(m)}_{\xi^{(2)}}\{\max_{T'\le n\le T'+\epsilon m^{1/5}} \gamma S_n - \gamma\cdot S_{T'} \le m^{1/3}\}$$
$$= P^{(m)}_{\xi^{(2)}}\{\max_{0\le n\le \epsilon m^{2/5}} \gamma\cdot S_n \le m^{1/3}\} \le P^{(m)}_{\xi^{(2)}}\{\gamma\cdot S_{[\epsilon m^{2/5}]} \le m^{1/3}\}$$
$$= E_{0,\theta^{(2)}(\xi^{(2)}_0)}\left\{\frac{dP^{(m)}_{\xi^{(2)}}}{dP_{(0,\theta^{(2)}_0)}}(S_1,\cdots,S_{[\epsilon m^{2/5}]}); \gamma\cdot S_{[\epsilon m^{2/5}]} \le m^{1/3}\right\}$$
$$\to \lim_{m\to\infty} P_{0,\theta^{(2)}(\xi^{(2)}_0)}\{\gamma\cdot S_{[\epsilon m^{2/5}]} \le m^{1/3}\}$$

by Lemma 1 and Sheffé's theorem.

It is clear that $P_{(0,\theta^{(2)}_0)}\{\gamma\cdot S_{[\epsilon m^{2/5}]} \le m^{1/3}\} \to 0$ as $m \to \infty$ by the strong law of large numbers.

The argument above shows that $|T - T'| = o_p(m^{1/2})$ which implies

$$\lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\left\{\frac{T - c/\mu^{(2)}_{\gamma}}{\sqrt{\left(1-\frac{\epsilon_0}{\mu^{(2)}_{\gamma}}\right)\sigma^2_{\xi^{(2)}}c(\mu^{(2)}_{\gamma})^{-3}}} < x, \gamma\cdot S_{T_c} - c < y\right\}$$
$$= \lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\left\{\frac{T' - c/\mu^{(2)}_{\gamma}}{\sqrt{\left(1-\frac{\epsilon_0}{\mu^{(2)}_{\gamma}}\right)\sigma^2_{\xi^{(2)}_0}c(\mu^{(2)}_{\gamma})^{-3}}} < x, \gamma\cdot S_{T_c} - c < y\right\}.$$

Now the idea is to restart the process from $S_{T'}$. Define $T'' = \inf\{n, \gamma\cdot S_{n+T'} > c\}$ and

$R_\delta = \{(v_1, \cdots, v_d); T'(1-\delta)\xi_i/m \le v \le T'(1+\delta)\xi_i/m,\ i = 1, \cdots, d\}$ where $\xi = (\xi_1, \cdots, \xi_d)$,

$$A_x = \left\{ \frac{u - c/\mu_\gamma^{(2)}}{\sqrt{\left(1 - \frac{c_0}{\mu_\gamma^{(2)}}\right) \sigma^2_{\xi_0^{(2)}} c(\mu_\gamma^{(2)})^{-3}}} < x \right\}$$

Then the probability above equals

$$\begin{aligned} &\lim_{m\to\infty} \int_{A_x} P^{(m-u)}_{\xi^{(2)}-v}\{\gamma S_{T''} - (c - \gamma \cdot S_{T'}) < y\} P^{(m)}_{\xi^{(2)}}\{T' \in du, S^{(2)}_{T'} \in dv\} \\ &= \lim_{m\to\infty} \int_{A_x \cap R_\delta} P^{(m-u)}_{\xi^{(2)}-v}\{\gamma S_{T''} - (c - \gamma \cdot S_{T'}) < y\} P^{(m)}_{\xi^{(2)}}\{T' \in du, S^{(2)}_{T'} \in dv\} \end{aligned}$$

since $\lim_{m\to\infty} P^{(m)}_{\xi}\{S_{T'} \in R_\delta\} = 1$.

Now by Theorem 2

$$|P^{(m-v)}_{\xi^{(2)}-v}\{\gamma \cdot S_{T'}c - (c - \gamma S_{T'}) < y\} - \lim_{c\to\infty} P_{(0,\theta_0^{(2)})}\{\gamma \cdot S_{T_c} - c < y\}| < \epsilon$$

on the set $A_x \cap R_\delta$ for δ sufficiently small. This implies if we first let $m \to \infty$ and then $\delta \to 0$. Then we have

$$\begin{aligned} &\lim_{m\to\infty} P^{(m)}_{\xi^{(2)}} \left\{ \frac{T - c/\mu_\gamma^{(2)}}{\sqrt{\left(1 - \frac{c_0}{\mu_\gamma^{(2)}}\right) \sigma^2_{\xi_0^{(2)}} c(\mu_\gamma^{(2)})^{-3}}} < x, \gamma \cdot S_{T_c} - c < y \right\} \\ &= \lim_{c\to\infty} P_{(0,\theta^{(2)}(\xi_0^{(2)}))}\{\gamma \cdot S_{T_c} - c < y\} \lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\{A_x\} \\ &= \Phi(x) \cdot \lim_{c\to\infty} P_{(0,\theta_0^{(2)})}\{\gamma \cdot S_{T_c} - c < y\} \end{aligned}$$

This completes the proof.

The next theorem is the main result of this chapter. It is also the result we need to complete the approximation in the previous section. Let $T_m = \inf\{n \ge m_0,\ mH(S_n/m) > 0\}$. From here to the end of the proof of Theorem 6 S_n will be defined as $S_n = \sum_{i=1}^n \tilde{X}_i$ where $\tilde{X}_i = (1, X_i)$. X_i $i = 1, 2, \cdots$ is an i.i.d. d-dimensional random vector sequence. The distribution of X_i can be imbedded in a d-parameter exponential family. The reason for defining S_n in this way is to include a more general stopping rule which is needed in applications. Before stating the theorem we list some assumptions which are easy to check when applying the theorem to a particular example, but it is tedious to impose conditions on the function H which imply these assumptions.

(I) $\exists t_0 > 0$ such that $\lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\left\{\left|\frac{T_m}{m} - t_0\right| > \epsilon\right\} = 0 \ \forall \epsilon > 0$.

(II) H is continuous in a cone C containing the line of drift $(1, \mu(0, \theta_0^{(2)}))$.

(III) H has continuous first partial in a neighborhood of $t_0\mu$ and $\gamma\cdot\mu \neq 0$ where $\gamma = \nabla H(t_0\mu)$.

Let $\mathcal{N} = \{(x_1, \cdots, x_{d+1}) : H(x_1, \cdots, x_{d+1}) < 0\}$. If assumptions (II) and (III) are true, then by the implicit function theorem there exists a continuously differentiable function $f(x_1, \cdots, x_d)$ satisfying $H(x_1, \cdots, x_d,\ g(x_1, \cdots, x_d)) = 0 \ \forall (x_1, \cdots, x_d)$ belongs to some neighborhood O of $(t_0\mu_1, \cdots, t_0\mu_d)$ and $g(t_0\mu_1, \cdots, t_0\mu_d) = t_0\mu_{d+1}$.

Let $\mathcal{F} = \{(x_1, \cdots, x_{d+1}) : x_{d+1} = g(x_1, \cdots, x_d)\}$ be a d-dimensional surface which is well defined on O and extend in some smooth way to R^d.

(IV) Boundary $(\mathcal{N}) \cap C = C \cap \mathcal{F}$.

Now we are ready to state

Theorem 6. If assumptions (I)–(IV) hold then

$$\lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\{mH(S_{T_m}/m) \le x\} = \lim_{c\to\infty} P_{(0,\theta_0^{(2)})}\{\gamma \cdot S_{T_c} - c \le x\}.$$

Before we go any further, let me make some remarks.

Remark 1: The problem is invariant under rotation so we may choose the coordinate system such that $\mu = (0, 0, \cdots, 0, \|\mu\|)$.

Remark 2: The limit on the right hand side of the equation above can be determined by the ordinary renewal theorem.

Although the proof of Theorem 3 is very complicated and technical, the basic strategy is not difficult to explain.

It will be shown that to determine $\lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\{mH(S_{T_m}/m) \le x\}$ is equivalent to finding the asymptotic distribution of excess over the hypersurface $m\mathcal{F}$ by the conditional random walk. In Lemma 4 below we shall exhibit a surface $m\mathcal{F}^{(k)}$ which is "near" $m\mathcal{F}$. The idea is to condition on the position of the random walk at the first time it crosses $m\mathcal{F}^{(k)}$ then restart the random walk from there. Lemmas 5 and 6 below amount to showing that the restarted random walk will cross $m\mathcal{J}$ in a time of magnitude $o(m^{1/2})$. Now we can simplify the problem in two directions. First, by Lemma 1 within the time of magnitude $o(m^{1/2})$ the conditional random walk behaves like an independent one. Secondly, since the restarted random walk will cross $m\mathcal{J}$ within time $o(m^{1/2})$ the only important part of $m\mathcal{J}$ is a set of diameters $o(m^{1/2})$, within which $m\mathcal{J}$ behaves like a hyperplane. (To every man on the earth, the earth is "flat"). Finally, an argument like the proof of Theorem 2 can be used to finish the proof. The proof of Theorem 3 is preceded by three lemmas.

Lemma 7. Let $m\mathcal{J}^{(k)} = m\mathcal{J} - k\mu/\|\mu\|$ where k is a positive number which may depend on m. There exist $\alpha_2 > \alpha_1 > 0$ such that $\alpha_2 k \geq d(m\mathcal{J} \cap C,\ m\mathcal{J}^{(k)} \cap C) \geq \alpha_1 k$.

Proof: Let $M = \sup_{x \in O} \|\nabla g(x)\|$, $\epsilon = (2M)^{-1}$. (If $M = 0$ then g is constant over O, and Lemma 1 holds trivially. In fact $d(m\mathcal{J} \cap C,\ m\mathcal{J}^{(k)} \cap C = k)$. Now

$$
\begin{aligned}
d(m\mathcal{J} \cap C,\ m\mathcal{J}^{(k)} \cap C)^2 &\geq \inf_{x,y \in \mathcal{J} \cap C} (\|m(x-y) - k\mu/\|\mu\|\|^2) \\
&= \inf_{x,y \in \mathcal{J} \cap C} (m^2\|x-y\|^2 - 2mk\langle x-y,\ \mu/\|\mu\|\rangle + k^2).
\end{aligned}
$$

Note that

$$
\begin{aligned}
m^2&\|x-y\|^2 - 2mk\langle x-y,\ \mu/\|\mu\|\rangle + k^2 \\
&= m^2[\|(x_1,\cdots,x_d) - (y_1,\cdots,y_d)\|^2 + |g(x_1,\cdots,x_d) - g(y_1,\cdots,y_d)|^2] \\
&\quad - 2mk[g(x_1,\cdots,x_d) - g(y_1,\cdots,y_d)] + k^2 \\
&= m^2\|(x_1,\cdots,x_d) - (y_1,\cdots,y_d)\|^2 + [m(g(x_1,\cdots,x_d) - g(y_1,\cdots,y_d)) - k]^2
\end{aligned}
\tag{4.2}
$$

where the first equality follows from $(\mu/\|\mu\|) \cdot x = g(x_1,\cdots,x_d)$. Let

$$
\epsilon_1 = \inf\{\|m(x-y) - k\mu/\|\mu\|\|;\ x,y \in \mathcal{J} \cap C,\ \|(x_1,\cdots,x_d) - (y_1,\cdots,y_d)\| > \epsilon k/m\}
$$

$$
\epsilon_1 = \inf\{\|m(x-y) - k\mu/\|\mu\|\|;\ x,y \in \mathcal{J} \cap C,\ \|(x_1,\cdots,x_d) - (y_1,\cdots,y_d)\| \leq \epsilon k/m\}.
$$

By (4.2) $\epsilon_1 \geq \epsilon^2 k^2$. By the mean value theorem

$$
|g(x_1,\cdots,x_d) - g(y_1,\cdots,y_d)| \leq M\|(x_1,\cdots,x_d) - (y_1,\cdots,y_d)\|
$$

so $\epsilon_2 \geq (k - \epsilon Mk)^2 \geq k^2/4$, clearly $d(m\mathcal{J} \cap C,\ m\mathcal{J}^{(k)} \cap C) \geq \epsilon_1 \wedge \epsilon_2$. Also $k \geq d(m\mathcal{J} \cap C,\ m\mathcal{J}^{(k)} \cap C)$ by the definition of $m\mathcal{J}^{(k)}$. This completes the proof of Lemma 1.

Let $x_1 = (x_1^{(1)}, \cdots, x_1^{(d)})$, $\|x_1\|_* = \sum_{i=1}^{d} |x_1^{(i)}|$. The following lemma is true.

Lemma 8. $\lim_{m\to\infty} mP_\xi^{(m)}\{\|X_1\|_* > m^\epsilon\} = 0 \ \forall \epsilon > 0$.

Proof: By sufficiency the conditional probability $P_\xi^{(m)}$ is independent of the parameter θ, so we may take $\theta = 0$, $E_0 X_1 = \xi_0$

$$mP_\xi^{(m)}\{\|X_1\|_* > m^\epsilon\} = m \int_{\|x\|_* > m^\epsilon} f(x) \cdot f_{m-1}(\xi - x)/f_m(\xi) dx. \tag{4.3}$$

By assumption $\sup_x |f_n(x)| \le B$ for all $n > n_0$, and proposition 1

$$f_m(\xi) \sim (2\pi m)^{-1/2} |\Sigma(\xi_0)|^{-1/2} \exp\{-n\phi(\xi_0)\}.$$

By (4.3) $\phi(\xi_0) = 0$, so for m sufficiently large

$$f_m(\xi) \ge (1/2)(2\pi m)^{-1/2} |\Sigma(\xi_0)|^{-1/2}.$$

Now

$$\int_{\|x\|_* > m^\epsilon} f(x) \cdot f_{m-1}(\xi - x)/f_m(\xi) dx \le 2B(2\pi m^3 |\Sigma(\xi_0)|)^{1/2} \int_{\|x\|_* > m^\epsilon} f(x) dx$$
$$\le 2B(2\pi m^3 |\Sigma(\xi_0)|)^{1/2} e^{-\lambda m^\epsilon} E(e^{\lambda \|X_1\|_*}) \quad \forall \lambda > 0.$$

Since the moment generating function of x_1 exists in a neighborhood of zero, $E(e^{\lambda\|X_1\|_*}) < \infty$ for λ sufficiently small.

It is clear that as $m \to \infty$

$$2B(2\pi m^3 |\Sigma(\xi)|)^{1/2} e^{-\lambda m^\epsilon} \cdot E(e^{\lambda \|X_1\|_*}) \to 0.$$

This completes the proof.

Define

$$T_m^{(k)} = \inf\{n \ge m_0 : S_n \text{ crosses } m\mathcal{F}^{(k)}\}.$$

Lemma 9. For any given $\delta > 0$ there exists m_δ such that for all $m > m_\delta$ the following inequality holds

$$P_{\xi^{(3)}}^{(m)}\{\alpha_4 m^{1/3} \ge d(S_{T_m^{(m^\epsilon/2)}}, m\mathcal{F} \cap C) \ge \alpha_3 m^{1/3}\} \ge 1 - \delta$$

where $\alpha_4 \ge \alpha_3 > 0$ are independent of m and δ.

Proof: Since with $P^{(m)}_{\xi(3)}$ probability close to one the conditional random walk S_n stays in the cone C for n sufficiently large, and we have

$$P^{(m)}_{\xi(3)}\{\|\tilde{X}_{T_m^{(m^{1/3})}}\|_* + m^{1/3} \geq d(S_{T_m^{(m^{1/3})}},\ m\mathcal{F} \cap C)\} \geq 1 - \frac{\delta}{2} \tag{4.4}$$

$$P^{(m)}_{\xi(3)}\{d(S_{T_m^{(m^{1/3})}}, m\mathcal{F} \cap C) \geq d(m\mathcal{F}^{(m^{1/3})} \cap C, m\mathcal{F} \cap C) - \|\tilde{X}_{T_m^{(m^{1/3})}}\|_*\} \geq 1 - \frac{\delta}{2}. \tag{4.5}$$

Clearly

$$|P^{(m)}_{\xi(3)}\{\|\tilde{X}_{T_m^{(m^{1/3})}}\|_* > m^\epsilon\} - P^{(m)}_{\xi(3)}\{\|\tilde{X}_{T_m^{(m^{1/3})}}\|_* > m^\epsilon,\ T_m^{(m^{1/3})} \leq \beta m\}| < \frac{\delta}{4}$$

for some constant $\beta > 0$ and m large enough and

$$\begin{aligned} &P^{(m)}_{\xi(3)}\{\|\tilde{X}_{T_m^{(m^{1/3})}}\|_* > m^\epsilon,\ T_m^{m^{1/3}} \leq \beta m\} \\ &\qquad \leq P^{(m)}_{\xi(3)}\{\max_{1\leq i\leq \beta m} \|X_i\|_* > m^\epsilon\} \leq \beta m P^{(m)}_{\xi(3)}\{\|X_1\|_* > m^\epsilon\}. \end{aligned}$$

The last quantity above tends to zero by Lemma 5. This implies

$$P^{(m)}_{\xi(3)}\{\|\tilde{X}_{T_m^{(m^{1/3})}}\|_* > m^\epsilon\} \leq \frac{\delta}{2}$$

for m sufficiently large. Choose $\epsilon < \frac{1}{3}$. In view of (4.4) and (4.5), the proof is completed.

Now we are ready to prove Theorem 3.

Proof of Theorem 3: Let $T_m^* = \inf\{n \geq m_0,\ S_n \text{ cross } m\mathcal{F}\}$. Since the conditional random walk S_n will stay in the cone C with high probability $\forall n \geq m_0$, by assumption (IV) $T_m^* = T_m$ with high probability. By Lemma 6 the conditional random walk starts at $S_{T_m^{(m^{1/3})}}$ will cross $m\mathcal{F}$ in a time that is $O_p(m^{1/3})$. Now the idea is to compare $m\mathcal{F}$ with its tangent plant at $mt_0\mu$ in a neighborhood with diameter that is $0(m^{1/3})$ via a Taylor expansion

$$mf\left(\frac{h}{m}\right) = mf(0) + \nabla f(0)' \cdot h + \frac{1}{m}h' \cdot \nabla^2 f\left(\delta\frac{h}{m}\right) \cdot h$$

where $\|h\| = 0(m^{1/3})$, $0 < \delta < 1$, so the conditional random walk starting at $S_{T_m^{(m^{1/3})}}$ crosses $m\mathcal{F}$ as if it crossing a hyperplane HP which is $0(m^{1/3})$ away,

$$HP = \{x;\ x \cdot \gamma = c_m\} \qquad c - m = 0(m^{1/3}).$$

S_n crosses HP when $S_n \cdot \gamma$ crosses the constant level C_m. The excess of interest turns out to be the excess in the normal direction, in the case of a hyperplane it equals $\gamma \cdot S_{T_{c_m}} - c_m$.

Let

$$A_1 = \{\text{Excess over } m\mathcal{J} \text{ in } \gamma = \nabla H(t_0\mu) \text{ direction } \leq x\}.$$

Then (by the arguments above)

$$\begin{aligned}\lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\{A_1\} &= \lim_{m\to\infty} E^{(m)}_{\xi^{(2)}}\{P^{(m)}_{\xi^{(2)}}[A_1 \mid \mathcal{E}_{T^{(m^{1/3})}_m}]\} \\ &= \lim_{m\to\infty} E^{(m)}_{\xi^{(2)}}\{P^{(m-T^{(m^{1/3})}_m)}_{\varsigma^{(2)}}[\gamma \cdot S_{T_{c_m}} - c_m \leq x]\}\end{aligned}$$

where $\mathcal{E}_{T^{(m^{1/3})}_m}$ is the σ-field generated by events prior to $T^{(m^{1/3})}_m$

$$\varsigma^{(2)} = \xi^{(2)} - S_{T^{(m^{1/3})}_m}.$$

Also let $\varsigma^{(2)}_0 = \xi^{(2)}/(m - T^{(m^{1/3})}_m)$ $A_2 = \{|\varsigma^{(2)}_0 - \xi^{(2)}_0| < \delta,\ T_{c_m} = O_p(m^{1/3})\}$. It is clear that A_2 has $P^{(m)}_{\xi^{(2)}}$ probability contain close to one as $m \to \infty$.

Now

$$\begin{aligned}&|E^{(m)}_{\xi^{(2)}}\{P^{(m-T^{(m^{1/3})}_m)}_{\varsigma^{(2)}}[\gamma S_{T_{c_m}} - c_m \leq x]\} - \lim_{c\to\infty} P_{(0,\theta^{(2)}(\xi^{(2)}_0))}[\gamma \cdot S_{T_c} - c \leq x]| \\ &\leq |E^{(m)}_{\xi^{(2)}}\{P^{(m-T^{(m^{1/3})}_m)}_{\varsigma^{(2)}}[\gamma \cdot S_{T_{c_m}} - c_m \leq x]\} - E^{(m)}_{\xi^{(2)}}\{P^{(m-T^{(m^{1/3})}_m)}_{\varsigma^{(2)}}(\gamma \cdot S_{T_{c_m}} - c_m \leq x);\ A_2\}| \\ &+ |E^{(m)}_{\xi^{(2)}}\{P^{(m-T^{(m^{1/3})}_m)}_{\varsigma^{(2)}}(\gamma - S_{T_{c_m}} - c_m \leq x);\ A_2\} \\ &\quad - E^{(m)}_{\xi^{(2)}}\{\lim_{c\to\infty} P_{(0,\theta^{(2)}(\varsigma^{(2)}_0))}(\gamma \cdot S_{T_c} - c \leq x);\ A_2\}| \\ &+ |E^{(m)}_{\xi^{(2)}}\{\lim_{c\to\infty} P_{(0,\theta^{(2)}(\varsigma^{(2)}_0))}(\gamma \cdot S_{T_c} - c \leq x); A_2\} \\ &- E^{(m)}_{\xi^{(2)}}\{\lim_{c\to\infty} P_{0,\theta^{(2)}(\xi^{(2)}_0))}(\gamma \cdot S_{T_c} - c \leq x); A_2\}| \\ &+ |E^{(m)}_{\xi^{(2)}}\{\lim_{c\to\infty} P_{(0,\theta^{(2)}(\xi^{(2)}_0)}(\gamma \cdot S_{T_c} - c \leq x);\ A_2\} - \lim_{c\to\infty} P_{(0,\theta^{(2)}(\xi^{(2)}_0))}(\gamma \cdot S_{T_c} - c \leq x)| \\ &= I_1 + I_2 + I_3 + I_4 \leq 4\epsilon\end{aligned}$$

where I_2 can be made arbitrarily small by Theorem 2 and $I_3 \leq \epsilon$ by Scheffé's theorem since the density of $P_{(0,\theta^{(2)}(\xi^{(2)}_0))}$ tends to that of $P_{(0,\theta^{(2)}(\xi^{(2)}))}$ as $m \to \infty$.

Let us define the excess R_{T_m} to be the quantity such that $S_{T_m} - \frac{\mu}{\|\mu\|} R_{T_m} \in m\mathcal{J}$. A Taylor expansion shows that

$$mH(S_{T_m}/m) = mH\left(\frac{S_{T_m} - (\mu/\|\mu\|)R_{T_m}}{m}\right) + \nabla H(x) \cdot \frac{\mu}{\|\mu\|} R_{T_m}.$$

The first term on the right hand side vanishes by the definition of $\mathcal{J}$.

The second term tends to the excess over $m\mathcal{J}$ in the normal direction of $m\mathcal{J}$ at μt_0 since $x \to \mu t_0$. We have

$$P^{(m)}_{\xi^{(2)}}\left\{mH\left(\frac{S_{T_m}}{m}\right) \leq x\right\} = \lim_{c\to\infty} P_{(0,\theta^{(2)}(\xi_0^{(2)}))}(\gamma \cdot S_{T_c} - c \leq x).$$

The proof is completed.

Combining the idea behind the proof of Theorem 5 and Theorem 6 it is not hard to see that a theorem like the following should be true. The proof is omitted, since every relevant step has appeared in the proofs of Theorems 5 and 6.

Theorem 10. Under the same assumption in Theorem 6, the following equality is true

$$\lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\left\{\frac{T_m - t_0 m}{\sqrt{(1-t_0)\sigma^2_{(0,\xi_0^{(2)})} t_0 m}} < x,\ mH(S_{T_m}/m) \leq y\right\}$$
$$= \Phi(x) \cdot \lim_{c\to\infty} P_{(0,\theta_0^{(2)})}\{\gamma \cdot S_{T_c} - c \leq y\}.$$

Now we are ready for some applications. The first application is the approximation of the conditional probability $P^{(m)}_{\xi}(T < m)$ in Chapter 3. By (3.5)

$$P^{(m)}_{\xi}(T < m) \sim (1 - t_{\xi_0})^{-d_1/2} |\Sigma[\xi_0)|^{1/2} \cdot |\Sigma[\xi_0^{(1)}/(1 - t_{\xi_0}),\ \xi_0^{(2)}]|^{-1/2}$$
$$E^{(m)}_{\xi^{(2)}}(e^{-R_m}) \exp\{-m[a_0 - \Lambda(\xi_0)]\}$$

where $R_m = (m-\tau)\Lambda[(\xi - S_\tau)/(m-\tau)] - a$ is the excess over the boundary at the stopping time $\tau = \inf\{n \leq m;\ (m-n)\Lambda[(\xi - S_n)/(m-n)] > a\}$. In this case $H(\mathbf{v}) = (1-t_0)\Lambda[(\xi_0 - v/1 - t_0)$ where $\mathbf{v} = (t_0, v) \in R^{d+1}$, $t_0 \in R$, $v \in R^d$. By Theorem 3

$$E^{(m)}_{\xi^{(2)}}(e^{-R_m}) \to \int_0^\infty e^{-x} P\{U_{\tau_+} > x\} dx \cdot [E(U_{\tau_+})]^{-1}$$

where $Y = \nabla H[t_{\xi_0}(1,\ \mu(0, \theta^{(2)}(\xi_0^{(2)})))] \cdot (1,\ X_1)$. X_1 is distributed according to $P_{(0,\theta^{(2)}(\xi^{(2)}))}$, and $U_n = \sum_{i=1}^n Y_i$.

The second application is concerned with the so-called "change point" problems. Assume that $X_1, X_2, \cdots, X_m$ are independent normally distributed random variables and that X_i has the mean value μ_i and variance 1. Suppose we are interested in testing the hypothesis $H_0 : \mu_1 = \cdots = \mu_m$ against the alternative H_1 : there exist k, $1 < k < m - 1$ such that $\mu_1 = \cdots = \mu_k \neq \mu_{k+1} = \mu_{k+2} = \cdots = \mu_m$, if k where known, the problem would be a two sample test of the equality of the means of the first sample $X_1, X_2, \cdots, X_k$ and second sample $X_{k+1}, X_{k+2}, \cdots, X_m$. For this problem the log likelihood ratio statitics would be $\Lambda_{k,m} = k(m-k)(\bar{X}_k - \bar{X}_{k,m})^2/m$ where $\bar{X}_k = k^{-1}\sum_1^k X_i$ and $\bar{X}_{k,m} - (m-k)^{-1}\sum_{k+1}^m X_i$. Since k is in fact unknown, the log likelihood ratio statistic is

$$\max_{1 \leq n \leq m-1} \Lambda_{n,m} \tag{4.6}$$

and the significance level of the likelihood ratio test is the probabiity under H_0 that the random variable (4.6) exceed some constant c.

Let $S_n = \sum_1^n X_i$. Simple algebra shows that

$$\Lambda_{n,m} = (S_n - nS_m/m)^2/n(1 - n/m).$$

It is easy to see that under H_0 the random varaibles

$$S_n - n\, S_m/m \qquad n = 1, 2, \cdots, m-1$$

have the same joint distribution as $S_1, S_2, \cdots, S_{m-1}$ given that $S_m = 0$ so the significance level is given by

$$P_0^{(m)}\{T \leq m-1\} \tag{4.7}$$

where

$$\begin{aligned} T &= \inf\{n : |S_n| \geq b[n(1-n/m)]^{1/2}\} \\ b &= c^{1/2}. \end{aligned} \tag{4.8}$$

We will prove an approximation formula which contains (4.7) as a special case.

Theorem 11. Let T be defined by (4.8), and assume that $b \to \infty$, $m_0 \to \infty$, $m_1 \to \infty$, and $m \to \infty$ in such a way that for some $0 < t_0 < t_1$ and $\mu_1 > 0$ $m_i/m \to t_i$ $(i = 0, 1)$,

$b/m^{1/2} = \mu_1$, then as $m \to \infty$

$$P_0^{(m)}\{m_0 < T \le m\} \sim 2(c/2\pi)^{1/2} e^{-c/2} \int_{\mu_1(t_1^{-1}-1)^{1/2}}^{\mu_1(t_0^{-1}-1)^{1/2}} \xi^{-1} v(\xi + \mu_1^2 \xi^{-1}) d\xi$$

where $v(\xi + \mu_1^2\xi^{-1})$ will be identified in the proof.

Proof: Let $Q^{(m)} = \int_{-\infty}^{\infty} P_\xi^{(m)} d\xi \cdot (2\pi)^{-1/2}$. An easy calculation shows that the likelihood ratio of $X_1, \cdots, X_n$ under $Q^{(m)}$ relative to $P_0^{(m)}$ is

$$[m(m-n)/n]^{1/2} \exp\left[S_n^2/2n(1-n/m)\right]$$

from which it follows by familiar argument that

$$\begin{aligned} P_0^{(m)}\{m_0 < T \le m_1\} &= m^{-1} \int_{-\infty}^{\infty} E_\xi^{(m)}\{[T/(1-T/m)]^{1/2} \exp\left[-(1/2)S_T^2/T(1-T/m)\right]; \\ &\quad m_0 < T \le m_1\} d\xi (2\pi)^{1/2} \\ &= \int_{-\infty}^{\infty} E_{m\xi}^{(m)} \left\{[T/(1-T/m)]^{1/2} \exp\left[-(1/2)S_T^2/T(1-T/m)\right];\ m_0 < T \le m_1\right\} d\xi \cdot (2\pi)^{-1/2} \\ &= (2\pi)^{-1/2} e^{-c/2} \int_{-\infty}^{\infty} E_{m\xi}^{(m)}\{[T/(1-T/m)]^{1/2} \\ &\quad \exp\left[-(1/2)(S_T^2/T(1-T/m) - c)\right];\ m_0 < T \le m\} d\xi. \end{aligned} \tag{4.9}$$

Solving the equation $\xi = \mu_1[t_\xi^{-1}(1-t_\xi)]^{1/2}$ for t_ξ, we have

$$t_\xi = \frac{1}{1+(\xi/\mu_1)^2}.$$

We know

$$P_{m\xi}^{(m)}\left\{\left|\frac{T}{m} - t_\xi\right| < \epsilon\right\} \to 1 \quad \forall \epsilon > 0 \tag{4.10}$$

Solving the following inequality for ξ, we have

$$t_1 < \frac{1}{1+(\xi/\mu_1)^2} < t_1 \Rightarrow \mu_1(t_1^{-1}-1)^{1/2} < \xi < \mu_1(t_0^{-1}-1)^{1/2} \tag{4.11}$$

or $-\mu_1(t_0^{-1}-1)^{1/2} < \xi < \cdot\ \mu_1(t_1^{-1}-1)^{1/2}$.

It is easy to see that

$$\lim_{m\to\infty} P_{m\xi}^{(m)}(m_0 < T \le m) = 1 \ \text{ for } \ \xi \ \text{ satisfying(4.11)}.$$

Applying Theorem 3 with $H(x,y) = (1/2)\left(\frac{x^2}{y(1-y)} - \mu_1^2\right)$ and $t_\xi\mu = (t_\xi\xi,\ t_\xi)$. we have

$$E_{m\xi}^{(m)}\left\{\exp\left[-(1/2)(S_T^2(1-T/m)-c)\right]\right\} \to \int_0^\infty [E(U_{\tau_+})]^{-1}e^{-x}P\{U_{\tau_+} > x\}dx$$

where

$$\begin{aligned} Y &= \nabla H(t_\xi\mu)\cdot(X_1+\xi,\ 1) \\ &= \left(\frac{\xi}{1-t_\xi}\right)\left(X_1 + \frac{\xi}{2(1-t_\xi)}\right) \\ &= \left(\xi + \frac{\mu_1^2}{\xi}\right)\left[X_1 + (1/2)\left(\xi + \frac{\mu_1^2}{\xi}\right)\right], \end{aligned}$$

$U_n = \sum_{i=1}^n Y_i$ and X_1 is distributed according to the standard normal.

Define

$$\int_0^\infty [E(U_{\tau_+})]^{-1}e^{-x}P\{U_{\tau_+} > x\} \equiv v(\xi + \mu_1^2\xi^{-1}). \tag{4.12}$$

Combining (4.10), (4.11), and (4.12) we obtain from (4.9) that

$$P_0^{(m)}\{m_0 < T \le m\} \sim 2(c/2\pi)^{1/2}e^{-c/2}\int_{\mu_1(t_1^{-1}-1)^{1/2}}^{\mu_1(t_0^{-1}-2)^{1/2}} \xi^{-1}v(\xi + \mu_1^2\xi^{-1})d\xi.$$

This completes the proof of Theorem 11.

Table 5.5

Significance Levels of Group Repeated t-Test

# of observations in a group	a	m_0	m	analytic approximation	Monte Carlo (2000 replications)*
2	3.65	8	40	0.050	0.052 ± 0.001
3	3.6	10	55	0.049	0.049 ± 0.001
4	3.6	10	70	0.051	0.052 ± 0.001
5	3.6	10	80	0.050	0.052 ± 0.001
7	3.6	15	120	0.047	0.047 ± 0.001

*Importance sampling is used in the Monte Carlo experiments above

Table 5.3

Significance Level of Modified Repeated t-Test

a	c	m_0	m	Analytic Approximation	Monte Carlo (6000 replications)
3.8	3.6	7	30	0.050	0.053 ± 0.001
3.95	3.6	7	40	0.050	0.052 ± 0.001
4.0	3.6	8	50	0.048	0.049 ± 0.0009
4.7	4.2	10	80	0.028	0.027 ± 0.0007
5.0	4.5	10	100	0.023	0.023 ± 0.0066

Table 5.4

Powers of Level of Modified Repeated t-Test

a	c	m_0	m	η	Analytic Approximation	Monte Carlo (2000 replications)
3.8	3.6	7	30	0.8	0.952	0.956 ± 0.005
3.95	3.6	7	40	0.7	0.960	0.959 ± 0.004
				0.5	0.717	0.727 ± 0.010
4.0	3.6	8	30	0.6	0.946	0.943 ± 0.005
				0.4	0.613	0.626 ± 0.011
4.7	4.2	10	80	0.5	0.947	0.937 ± 0.006
				0.4	0.779	0.770 ± 0.010
5.0	4.5	10	100	0.45	0.940	0.938 ± 0.005
				0.3	0.553	0.550 ± 0.001

Table 5.1

Significance Level of Repeated t-Test

a	m_0	m	Analytic Approximation	Monte Carlo (2000 replications)
3.8	7	30	0.052	0.053 ± 0.001
4.0	8	50	0.047	0.048 ± 0.001
4.5	10	75	0.032	0.033 ± 0.0006
5.0	10	110	0.024	0.023 ± 0.0004

*Importance sampling is used in the Monte Carlo experiments above

Table 5.2

Powers of Repeated t-Test

a	m_0	m	η	Analytic Approximation	Monte Carlo (2000 replications)
3.8	7	30	0.8	0.946	0.951 ± 0.005
			0.6	0.734	0.742 ± 0.010
4.0	8	30	0.6	0.934	0.933 ± 0.006
			0.4	0.584	0.596 ± 0.008
4.0	10	75	0.5	0.950	0.948 ± 0.005
			0.3	0.518	0.522 ± 0.011
5.0	10	110	0.4	0.882	0.889 ± 0.007
			0.3	0.581	0.581 ± 0.011

where in the equality above we have used Theorem 2 of Chapter 2 to obtain

$$\begin{aligned}\nu_-(\hat{\xi}_1,\hat{\xi}_2) &= \nu_+\{[\exp(2a_0/(1-\hat{t}_{\xi_0}))-1]^{1/2}\}\cdot\{(1/2)\log(1+\hat{\theta}^2)/[\hat{\theta}^2-(1/2)\log(1+\hat{\theta}^2)]\} \\ &= \nu_+\{[\exp(2a_0/(1-\hat{t}_{\xi_0}))-1]^{1/2}\}\cdot[a_0/(1-\hat{t}_{\xi_0})] \\ &\quad\cdot\{\exp[2a_0/(1-\hat{t}_{\xi_0})]-1-a_0/(1-\hat{t}_{\xi_0})\}^{-1}\end{aligned}$$

and the equality

$$\begin{aligned}|\ddagger(\hat{\xi}_1(1-\hat{t}_{\xi_0})^{-1},\hat{\xi}_2)|^{-1/2} &= 2^{-1/2}\cdot[\hat{\xi}_2-\hat{\xi}_1^2(1-\hat{t}_{\xi_0})^{-2}]^{-3/2} \\ &= 2^{-1/2}\cdot\hat{\xi}_2^{-3/2}\cdot\exp[3a_0/(1-\hat{t}_{\xi_0})]\end{aligned}$$

Next we consider the group repeated t-test. The stopping rule we are interested in here is

$$\tau_k = \inf\left\{n;\ n=m_0+ik, i=0,\cdots,\left[\frac{m-m_0}{k}\right], n\Lambda(S_n/n)>a\right\}$$

where k is the number of observations in a group. It is easy to see that τ_k is a stopping time. A moment's reflection we find that the corresponding significance levels and powers are the same as $k=1$, except the excess over the boundary part. To find the excess over the boundary part, it is sufficient to identify the increment of the random walk which generates the excess over the boundary. In this case, using the forward method, the corresponding increment of the random walk is $U_k=\sum_{i=1}^k Y_i$ where the distribution of Y_i's are given in (5.2). Using the backward method, the increment of interest is $V_k=\sum_{i=1}^k Z_i$, where the distribution of the Z_i's is given by (5.4).

Tables 5.1 – 5.4 below give some examples of the approximation of powers and significance levels of both RST and MRST. For comparison, the results of Monte Carlo experiments are also included. For details of numerical computation, see the Appendix.

What remains is to approximate

$$\int\int_A e^{-m(\phi_0(\xi_2)-\eta\xi_1)}d\xi_2 d\xi_1.$$

Since $(\hat{\xi}_1, \hat{\xi}_2)$ is located at the boundary of the set A, the multidimensional Laplace method does not apply. The argument below uses a chance of variable to convert the integral into a form which can be handled by iterative approximation. The change of variable is suggested by Fig. 1.

$$\begin{aligned}
&\int\int_A \exp[-m(\phi_0(\xi_2) - \eta\xi_1)]d\xi_2 d\xi_1 \\
&\quad = \int_0^\infty \int_{C_1}^{C_0} \exp[-m(\phi_0(y_2) - \eta y_1 y_2^{1/2}]y_2^{1/2} dy_2 dy_1 \quad \text{where } y_1 = \xi_1\xi_2^{-1/2} \; y_2 = \xi_2 \\
&\quad = \int_0^\infty \exp[-m(\phi_0(y_2))]y_2^{1/2} \left[\int_{C_1}^{C_0} \exp(m\eta y_1 y_2^{1/2})dy_1\right] dy_2 \\
&\quad = \int_0^\infty \exp[-m(\phi_0(y_2))]y_2^{1/2}\{(m\eta y_2^{1/2})^{-1}[\exp(m\eta_2^{1/2}c_0) - \exp(m\eta y_2^{1/2}c_1)]\}dy_2 \\
&\quad \sim \int_0^\infty \exp[-m(\phi_0(y_2))] \cdot (m\eta)^{-1} \cdot \exp(m\eta y_2^{1/2}c_0)dy_2 \\
&\quad = (m\eta)^{-1} \int_0^\infty e^{-mg(y_2)}dy_2
\end{aligned} \tag{5.23}$$

where the function $g(\cdot)$ is defined by (5.19). Now Laplace's method can be used to approximate

$$(m\eta)^{-1} \int_0^\infty e^{-mg(y_2)}dy_2 \sim (m\eta)^{-1}[2\pi/(mg''(\hat{\xi}))]^{1/2}e^{mg(\hat{\xi}_2)}.$$

By (5.22), (5.23)

$$\begin{aligned}
&P_\eta\{m_0 \le \tau < m, \Lambda(S_m/m) < c/m\} \sim \exp[-m(a_0 + \eta^2/2 + g(\hat{\xi}_2)] \\
&\quad \cdot \nu_-(\hat{\xi}_1, \hat{\xi}_2) \cdot \eta^{-1}[2\pi m g''(\hat{\xi}_2) \cdot (1 - \hat{t}_{\xi_0}) \cdot |\mathfrak{T}(\hat{\xi}_1(1 - \hat{t}_{\xi_0})^{-1}, \hat{\xi}_2)|]^{1/2} \\
&\quad = \exp[-m(a_0 + \eta^2/2 + g(\hat{\xi}_2)] \cdot \nu_+\{[\exp[2a_0(1 - \hat{t}_{\xi_0})^{-1}] - 1]^{1/2}\} \cdot \exp[3a_0/(1 - \hat{t}_{\xi_0})] \cdot a_0 \\
&\quad \eta^{-1} \cdot \{\exp[2a_0/(1 - \hat{t}_{\xi_0})] - a_0(1 - \hat{t}_{\xi_0})^{-1}\}^{-1} \cdot [2\pi m(\hat{\xi}_2 + (1/2)c_0\eta\hat{\xi}_2^{3/2})(1 - \hat{t}_{\xi_0})^3]^{1/2}
\end{aligned} \tag{5.24}$$

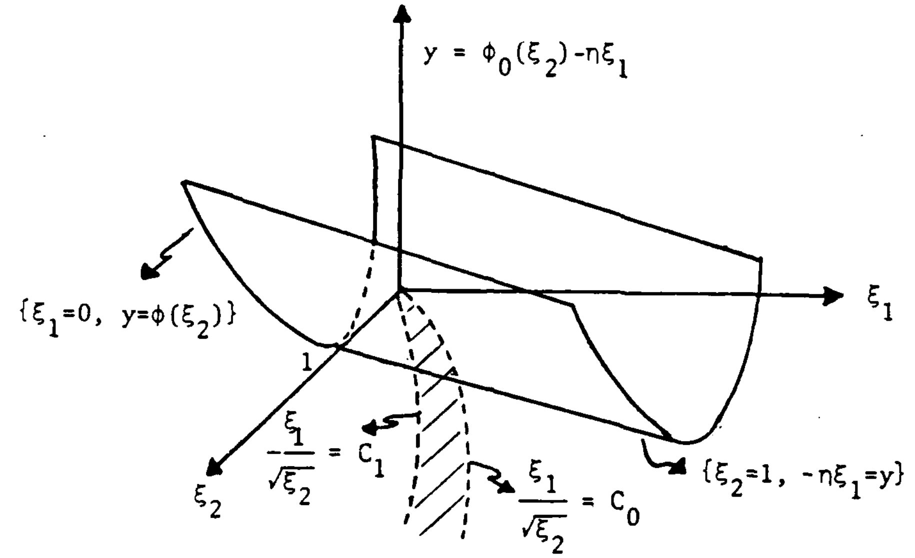


Figure 1

Since $g''(\xi_2) > 0$ $g(\xi_2)$ has minimum.

Let $\xi_2^{-1/2} = x$ be the minimum satisfying

$$x^2 + c_0\eta x - 1 = 0 \Rightarrow x = \{[(c_0\eta)^2 + 4]^{1/2} - c_0\eta\}/2$$

$$\hat{\xi}_2 = x^{-2}\{2/[(c_0^2\eta^2 + 4)^{1/2} - c_0\eta]\}^2 = [(c_0^2\eta^2 + 4)^{1/2} + c_0\eta]^2/4 \tag{5.20}$$

$$\hat{\xi}_1 = c_0\hat{\xi}_2^{1/2} = c_0 x^{-1} = \frac{c_0}{2}(c_0^2\eta^2 + 4)^{1/2} + c_0\eta. \tag{5.21}$$

As we pointed out earlier the only important part of the integral (5.17) is the integral over an arbitrarily small neighborhood N of $(\hat{\xi}_1, \hat{\xi}_2)$. Part of the intergrand of (5.17) is approximately constant over N, namely

$$\begin{aligned}&(1 - t_{\xi_0})^{-1/2}|\Sigma[\xi_1(1 - t_{\xi_0})^{-1}, \xi_2]|^{-1/2}\nu_-(\xi_0)\\ &\qquad\approx (1 - \hat{t}_{\xi_0})^{-1/2}|\Sigma[\hat{\xi}_1(1 - \hat{t}_{\xi_0})^{-1}, \hat{\xi}_2]|^{-1/2}\nu_-(\hat{\xi}_0) \text{ over } N\end{aligned} \tag{5.22}$$

where $\hat{t}_{\xi_0}$ satisfies $(-1/2)(1 - \hat{t}_{\xi_0})\log\{1 - [\hat{\xi}_1 \cdot \hat{\xi}_2^{-1/2}(1 - [\hat{t}_{\xi_0})^{-1}]^2\} = a/m$.

$\phi_0(\xi_2)$ in a neighborhood of 1

$$\begin{aligned}\phi_0(\xi_2) &= \phi_0(1) + \phi'(1)(\xi_2 - 1) + (1/2)\phi''(*)(\xi_2 - 1)^2 \\ &\sim (1/4)(\xi_2 - 1)^2\end{aligned},$$

so

$$\int_{I_m} e^{-m\phi_0(\xi_2)} d\xi_2 \sim \int_{-\infty}^{\infty} \exp(-mx^2/4)dx = (4\pi/m)^{1/2}. \tag{5.16}$$

Substituting (5.16) into (5.15) we have

$$\begin{aligned}&P_0\{m_0 \le \tau < m,\ \Lambda(S_m/m) < c/m\} \\ &\sim e^{-a}(m/\pi)^{1/2} \int_{A|\xi_2=1} (1 - t_{\xi_0})^{-1/2} |\Sigma(\xi_1(1 - t_{\xi_0})^{-1}, 1)|^{-1/2},\ \nu_-(\xi_1, 1)d\xi_1.\end{aligned}$$

The change of variable $\log(1 + \theta^2) = -\log[1 - \xi_1^2(1 - t_{\xi_0})^{-2}]$ transforms the expression above into $2e^{-a}(a/\pi)^{1/2}\int_{\bar{\theta}}^{\theta_0}[\log(1 + \theta^2)]^{-1/2}\nu_+(\theta)d\theta$ in agreement with (5.1).

Now let us bring our attention back to (5.14) with $\eta \neq 0$. The integral to be evaluated is

$$\int\int_A (1 - t_{\xi_0})^{-1/2} |\Sigma[\xi_1(1 - t_{\xi_0})^{-1}, \xi_2]|^{-1/2} \nu_-(\xi_0) \exp[-m(\phi_0(\xi_2) - \eta\xi_1)] d\xi_2 d\xi_1. \tag{5.17}$$

The set

$$\begin{aligned}A &= \{0 < t_{\xi_0} \le 1 - m_0/m,\ \Lambda(\xi_0) < c/m\} \\ &= \{m_0/m \le t_1 < 1,\ (1/2)\log\{1/[1 - \xi_1/\xi_1^{1/2}]\} < c/m\} \\ &= \{c_1 < \xi_1/\xi_2^{1/2} < c_0\}\end{aligned} \tag{5.18}$$

where $t_1 = 1 - t_{\xi_0}$ and $c_0 = (1 - e^{-\frac{2c}{m}})^{1/2}$. c_1 satisfies $(m_0/2m)\log\{1/[1 - (c_1 m_0)^2/m]\} = a_0$. It is clear that the only part which is of first order importance in (5.18) is the integral over a small neighborhood of $(\hat{\xi}_1, \hat{\xi}_2)$ where $(\hat{\xi}_1, \hat{\xi}_2)$ minimize $\phi_0(\xi_2) - \eta\xi_1$ over A.

We now proceed to identify $(\hat{\xi}_1, \hat{\xi}_2)$. The following picture helps us locate $(\hat{\xi}_1, \hat{\xi}_2)$.

It is clear from Fig. 1 that the minimum of $\phi(\xi_2) - \eta\xi_1$ over A occurs on the curve $\xi_1/\xi_2^{1/2} = c_0$. On this curve $\phi_0(\xi_2) - \eta\xi_1$ is equal to

$$(1/2)(\xi_2 - 1 - \log\xi_2) - \eta c_0 \xi_2^{1/2} = g(\xi_2). \tag{5.19}$$

$$\frac{dg(\xi_2)}{d\xi_2} = (1/2)(1 - \xi_2^{-1} - c_0\eta\xi_2^{-1/2}).$$

Next we consider the problem of approximating the power of the modified repeated T-test. It is easy to see that the powers of the modified test depend on the parameter μ, σ^2 only through $\eta = \mu/\sigma$, so without loss of generality we may take $\sigma = 1$ and $\mu = \eta$. By symmetry the power at η equals that at $-\eta$, so we may assume $\eta > 0$. The power of this test at η, by definition, is equal to

$$\begin{aligned} &P_\eta\{m_0 \le \tau \le m\} + P_\eta\{\tau > m,\ \Lambda(S_m/m) \ge c/m\} \\ &\qquad = P_\eta\{m_0 \le \tau < m,\ \Lambda(S_m/m) < c/m\} + P_\eta\{\Lambda(S_m/m) \ge c/m\}. \end{aligned} \tag{5.13}$$

The second part on the right hand side of (5.13) can be easily obtained by approximating the tail probabiities of the noncentral T-distribution, so it is sufficient to approximate $P_\eta\{m_0 \le \tau < m,\ \Lambda(S_m/m) < c/m\}$. By Proposition 1 of Chapter 2 under P_η, S_m has asymptotic density in the following form,

$$f_{m,\eta(m\xi)} \sim (z\pi m)^{-1}|\Sigma(\xi)|^{-1/2}\exp[-m(\phi(\xi_0) - \eta\xi_1 + \eta^2/2)].$$

Unconditioning (5.2) with respect to $f_{m,\eta}(m\xi)$ gives

$$\begin{aligned} e^{-a}(m/2\pi)\int_{\tilde{A}}\int_{A|\xi_1} &(1-t_{\xi_0})^{-1/2}|\Sigma[\xi_1(1-t_{\xi_0})^{-1}, \xi_2]|^{-1/2}\nu_-(\xi_0) \\ &\cdot \exp[-m(\phi_0(\xi_2) - \eta\xi_1 + \eta^2/2)]d\xi_2 d\xi_1 \end{aligned} \tag{5.14}$$

where $A = \{0 < t_0 < 1 - m_0/m,\ \Lambda(\xi_0) < c/m\}$, $A_{|\xi_1}$ denotes the cross section of A in the ξ_2 direction when ξ_1 is given and $\tilde{A} = \{\xi_1,\ 0 < t_0 < 1 - m_0/m,\ \Lambda(\xi_0) < c/m\}$. Observe that when $\eta = 0$ (5.14) should reduce to (5.1). The following argument shows that indeed it does.

When $\eta = 0$ the integral part of (5.14) reduces to

$$\int_{\tilde{A}}\left[\int_{A|\xi_1}(1-t_{\xi_0})^{-1/2}|\Sigma(\xi_1(1-t_{\xi_0})^{-1}, \xi_2)|^{-1/2}\nu_-(\xi_0)e^{-m\phi_0(\xi_2)}d\xi_2\right]d\xi_1. \tag{5.15}$$

Since $\phi_0(1) = 0 = \min_x \phi_0(x)$, it is easy to see thta the integral over the interval $I_m = (1-\epsilon_m, 1+\epsilon_m)$ with $\lim_{m\to\infty}\epsilon_m = 0$ constitutes the major contribution of the inner integral in (5.13). For m sufficiently large $(1-t_{\xi_0})^{-1/2}|\Sigma(\xi_1(1-t_{\xi_0})^{-1}, \xi_2)|^{-1/2}\nu_-(\xi_0)$ is effectively constant (with respect to ξ_2) over I_m. We still have to evaluate $\int_{I_m} e^{-m\phi_0(\xi_2)}d\xi_2$. Expanding

The first term on the right hand side of (5.10) can be computed easily by calculating the tail probability of the t-distribution. So the main task here is to approximate the second term $P_0\{\tau \le m,\ \Lambda(S_m/m) \le c/m\}$. Using the backward method, the procedure of approximating $P_0\{\tau \le m,\ \Lambda(S_m/m) < c/m\}$ is no more complicated than that of $P_0\{\tau \le m\}$. The only difference is the range of integration is changed. To determine the range of integration, let us recall that in obtaining (5.8), we made the following change of variable $\log 1 + \theta^2 = -\log[1 - (z/1 - t_{\xi_0})^2]$, where $z = \xi_1 \xi_2^{-1/2}$, z and t_{ξ_0} satisfying $(1 - t_{\xi_0}) \log\{[1 - (z(1 - t_{\xi_0})^{-1}]^{-1}\} = 2a_0$. Now

$$\{\xi : \Lambda(\xi) < c/m\} \Leftrightarrow \{z : (1/2)\log[(1 - z^2)^{-1}] < c/m\} \Leftrightarrow \{z : |z| < \theta_2^2(1 + \theta_2^2)^{-1}\}$$

where θ_2 satisfies $c/m = (1/2)\log(1 + \theta_2^2)$. Clearly

$$1 - t_{\xi_0} = 2a_0[\log(1 + \theta^2)]^{-1} = \log(1 + \theta_1^2)[\log(1 + \theta^2)]^{-1}$$

θ_1 satisfies $a/m = (1/2)\log(1 + \theta_1^2)$. Substituting $z = \theta_2^2(1 + \theta_2^2)^{-1}$ and $1 - t_{\xi_0} = \log(1 + \theta_1^2) \cdot [\log(1 + \theta^2)]^{-1}$ into $\log(1 + \theta^2) = \log\{[1 - z^2(1 - t_{\xi_0})^{-2}]^{-1}\}$ we find the lower bound of the range of integration $\tilde{\theta}$ satisfies

$$\frac{\tilde{\theta}^2}{1 + \tilde{\theta}^2}\left[\frac{\log(1 + \theta_1^2)}{\log(1 + \tilde{\theta}^2)}\right] = \frac{\theta_2^2}{1 + \theta_2^2}. \tag{5.11}$$

Of course we take only the root $\tilde{\theta} > 0$ of (5.11). Note that when $\theta_2 = \theta_1$, $\tilde{\theta} = \theta_1$. The upper bound of the range of integration remains the same. We have

$$P_0\{m_0 \le \tau < m,\ \Lambda(S_m/m) < c/m\} \sim 2(a/\pi)^{1/2} e^{-a} \int_{\tilde{\theta}}^{\theta_0} [\log(1 + \theta^2)]^{-1/2} \nu_+(\theta) d\theta. \tag{5.12}$$

The sharp-eyed reader may discover the possibility of obtaining (5.12) by modifying Siegmund's method, mentioned in the last paragraph of page 32, along the same line of arguments on pages 11–13. This possibiity does not exist because there exists a deep-buried measurability problem. To make a long story short, the two dimensional sufficient process $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ is not measurable with respect to the σ-field generated by the maximum invariant process $(X_1^{-1}X_2, X_1^{-1}X_2, \cdots, X_1^{-1}X_n)$.

so Y has density

$$f_Y(y) = 2(1+\theta^2)\theta^{-2}\sum_{j=0}^{\infty} e^{-1/2\theta^2}(2\theta^2)^{-j}(j!)^{-1}[\Gamma(j+1/2)]^{-1}2^{-(j+1/2)}$$
$$\cdot \exp\{(1+\theta^2)\cdot[y-1/2-(1/2)\log(1+\theta^2)]\theta^{-2}\}$$
$$\cdot\{\theta^2(1+\theta^2)^{-1}[1+\log(1+\theta^2)-2y]^{j-1/2}\cdot 1_{(-\infty,1/2+(1/2)\log(1+\theta^2))}(y).$$

After some simple algebra we find that

$$f_Y(y) = \theta^{-1}\exp\{(1+\theta^2)y-(2\theta^2)^{-1}[2+\theta^2+\log(1+\theta^2)]\}\cdot\sum_{j=0}^{\infty}(\theta^{-4}+\theta^{-2})^j$$
$$\cdot[\Gamma(j+1/2)]^{-1}\cdot 2^{-j}\cdot(j!)^{-1}\cdot\{1/2+(1/2)\log(1+\theta^2)-y\}^{j-1/2}\cdot 1_{(-\infty,1/2+(1/2)\log(1+\theta^2))}$$

$-Z$ has density

$$f_{-Z}(y) = 2\cdot\theta^{-2}\sum_{j=0}^{\infty}\exp[-(2\theta^2)^{-1}(1+\theta^2)][(2\theta^2)^{-1}(1+\theta^2)]^j(j!)^{-1}[\Gamma(j+1/2)]^{-1}\cdot 2^{-(j+1/2)}$$
$$\cdot\exp\{\theta^{-2}[y-(1/2)-(1/2)\log(1+\theta^2)]\}\cdot\{\theta^{-2}[1+\log(1+\theta^2)-2y]\}^{j-1/2}$$
$$\cdot 1_{(-\infty,1/2[1+\log(1+\theta^2)])}(y)$$
$$=\theta^{-1}\exp\{\theta^{-2}[y-1-\theta^2/2-(1/2)\log(1+\theta^2)]\}\cdot\sum_{j=0}^{\infty}(\theta^{-4}+\theta^{-2})^j\cdot(j!)^{-1}$$
$$\cdot[\Gamma(j+1/2)]^{-1}2^{-j}\{(1/2)[1+\log(1+\theta^2)]-y\}^{j-1/2}\cdot 1_{(-\infty,1/2[1+\log(1+\theta^2)])}(y).$$

The likelihood ratio of Y with respect to $-Z$ is surprisingly simple $f_Y(x)/f_{-Z}(x)=e^x$. Now applying Theorem 2 of Chapter 2 we have

$$\nu_+(\theta)=\nu_-(\theta)\mu_Z/\mu_Y. \tag{5.9}$$

Simple algebra shows that $\mu_Z = EZ = \theta^2-(1/2)\log(1+\theta^2)$, $\mu_Y = EY = (1/2)\log(1+\theta^2)$. Now by (5.9) it is clear that (5.8) and (5.1) agree. Next we consider the modified repeated t-test. The stopping rule is still τ, but we add the set $\{m\Lambda(S_m/m)>c\}$ to the rejection region, that is, we reject the null hypothesis when either $\tau\le m$ or $\{\tau<m,\ m\Lambda(S_m/m)>c\}$, where $0<a$. The significance elvel of this test is

$$P_0\{\tau\le m\}+P_0\{\tau>m,\ \Lambda(S_m/m)>c/m\}=P_0\{(S_m/m)>c/m\}$$
$$+P_0\{\tau\le m,\ \Lambda(S_m/m)\le c/m\}. \tag{5.10}$$

$$\frac{dy}{dz} = (m-1)^{1/2}\frac{d}{dz}[z(1-z^2)^{-1/2}] = (m-1)^{1/2}(1-z^2)^{-3/2} = (m-1)^{1/2}[\xi_2/(\xi_2-\xi_1^2)]^{3/2}$$

and $\xi_2^{3/2}\cdot[\xi_2-\xi_1^2(1-t_{\xi_0})^{-2}]^{-3/2} = \exp[3a_0(1-t_{\xi_0})^{-1}]$. Substituting these results into the integral above and using Stirling's formula on the gamma functions we have

$$e^{-a}(m/2\pi)^{1/2}\cdot\int_{\{0<t_{\xi_0}<1-m_0/m\}}\exp[3a_0(1-t_{\xi_0})^{-1}](1-t_{\xi_0})^{-1/2}\nu_-[\xi_0(z)]dz. \tag{5.6}$$

We need to make another change of variable. Let

$$\log(1+\theta^2) = -\log[1-(z/1-t_{\xi_0})^2] = 2a_0(1-t_{\xi_0})^{-1}. \tag{5.7}$$

Observe that $t_{\xi_0}=0$ implies $\theta=(e^{2a/m}-1)^{1/2}$ and $t_{\xi_0}=1-m_0/m \Rightarrow \theta=(e^{2a/m_0}-1)^{1/2}$, $\frac{dz}{d\theta}=(1-t_{\xi_0})(1+\theta^2)^{-3/2}[\log(1+\theta^2)-2\theta^2][\log(1+\theta^2)]^{-1}$. Substituting these results into (5.4) gives

$$\begin{aligned} P_0\{m_0\le\tau<m\} \sim 2(a/\pi)^{1/2}e^{-a}\int_{\theta_1}^{\theta_0}&[\log(1+\theta^2)]^{-1/2} \\ &\nu_-[\xi_0(\theta)][2\theta^2-\log(1+\theta^2)]\cdot[\log(1+\theta^2)]^{-1}d\theta \end{aligned} \tag{5.8}$$

where $\nu_-[\xi_0(\theta)] = \lim_{b\to\infty}E\{e^{-(V_{\tau_b}-b)}\}$ where $V_n=\sum_{i=1}^n Z_i$, $Z_1,Z_2,\cdots$ is an i.i.d. sequence of random variables, each Z_i has the same distribution as

$$Z=(\theta^2/2)X^2-\theta(1+\theta^2)^{1/2}X+(1/2)[\theta^2-\log(1+\theta^2)], \quad X\sim N(0,1).$$

To show that (5.6) agrees with (5.1) we write Y,Z in the following form

$$Y=-(1/2)\theta^2(1+\theta^2)^{-1}(x-\theta^{-1})^2+(1/2)(1+\log(1+\theta^2))$$
$$Z=(\theta^2/2)(X-(1+\theta^2)^{1/2}\theta^{-1})^2-(1/2)[1+\log(1+\theta^2)]$$

Let $\chi_1^2(\gamma)$ denote the noncentral χ^2-distribution with one degree of freedom and noncentral parameter γ. It is easy to see that Y, and $-Z$ distributed as $-(1/2)[\theta^2/(1+\theta^2)]\chi_1^2(\theta^{-2})+(1/2)(1+\log(1+\theta^2))$ and $-(1/2)\theta^2\chi_1^2(\theta^{-2}+1)+(1/2)[1+\log(1+\theta^2)]$ respectively, and they have the same support $(-\infty,\ (1/2)[1+\log(1+\theta^2)])$, hence the likelihood ratio of Y with respect to $-Z$ exists. $\chi_1^2(\gamma)$ has density of the following form (see e.g. Ferguson (1967))

$$\sum_{j=0}^{\infty}e^{-\gamma/2}(\gamma/2)^j(j!)^{-1}[\Gamma(j+1/2)]^{-1}2^{-(j+1/2)}\cdot e^{-x/2}x^{j-1/2}\cdot 1_{(0,\infty)}(x),$$

x); $(\xi_2 - x)/(1-x)$]. Straightforward calculation gives

$$
\begin{aligned}
Z &= \nabla H(t_{\xi_0}\mu)\cdot(1,\ \sqrt{\xi_2}X,\ \xi_2 X^2) \\
&= (1/2)[\xi_1/(1-t_{\xi_0})]^2\{\xi_2 - [\xi_1/(1-t_{\xi_0})]^2\}^{-1}X^2 - \{\xi_1/[\xi_2^{1/2}\cdot(1-t_{\xi_0})]\} \\
&\quad\cdot\{\xi_2 - [\xi_1/(1-t_{\xi_0})]^2\}^{-1}X + (1/2)\{\xi_1^2(1-t_{\xi_0})^{-2} \\
&\quad\cdot[\xi_2 - \xi_1^2(1-t_{\xi_0})^{-2}]\}^{-1} - \log\{\xi_2[\xi_2 - \xi_1^2(1-t_{\xi_0})^{-2}]^{-1}\}\}.
\end{aligned}
\tag{5.4}
$$

where X is distributed according to the standard normal. In order to obtain the significance levels from (5.3) we have to uncondition (5.3) with respect to $P_0\{S_m \in dm\xi_0\}$. Observe that each term on the right hand side of (5.3) is a function of $z = \xi_1\xi_2^{1/2}$, hence a function of $y = (m-1)^{1/2}z\cdot(1-z^2)^{-1/2}$. This reduces the conditional probability $P_\xi^{(m)}$ which in general is a function of two variables to a function of one variable. This is because the likelihood ratio statistic is invariant under scale change. It is easy to see that y has a t-distribution with $(m-1)$ degrees of freedom. Now multiplying (5.3) by the density of y and integrating over the appropriate range give

$$
\begin{aligned}
e^{-a}\int_{\{0<t_{\xi_0}<1-m_0/m\}} &(\xi_2-\xi_1^2)^{3/2}\cdot[\xi_2-\xi_1^2(1-t_{\xi_0})^{-2}]^{-3/2} \\
&(1-t_{\xi_0})^{-1/2}\nu_-(\xi_0)e^{m\Lambda(\xi_0)}g_{m-1}(y)dy
\end{aligned}
\tag{5.5}
$$

where $g_{m-1}(y) = \Gamma(m/2)[(m-1)\pi]^{-1/2}\cdot[\Gamma\{(m-1)/2)\}]^{-1}\cdot[y^2/(m-1)+1]^{-m/2}$ is the density of the t-distribution with $(m-1)$ degrees of freedom.

The second factor on the right hand side of the equation above cancels with

$$
e^{m\Lambda(\xi)} = \exp\{(m/2)\log[\xi_2/(\xi_2-\xi_1^2)]\} = [\xi_2/(\xi_2-\xi_1^2)]^{m/2} = [y^2/(m-1)+1]^{m/2}
$$

and (5.5) reduces to

$$
\begin{aligned}
&e^{-a}\Gamma(m/2)[(m-1)\pi]^{-1/2}[\Gamma\{(m-1)/2\})]^{-1}\int_{\{0<t_{\xi_0}\le 1-m_0/m\}}(\xi_2-\xi_1^2)^{3/2} \\
&\quad[\xi_2-\xi_1^2(1-t_{\xi_0})^{-2}]^{-3/2}\cdot(1-t_{\xi_0})^{-1/2}\nu_-(\xi_0)dy \\
&= e^{-a}\Gamma(m/2)[(m-1)\pi]^{-1/2}[\Gamma\{(m-1)/2\}]^{-1}\int_{\{<t_{\xi_0}\le 1-m_0/m\}} \\
&\quad(\xi_2-\xi_1^2)^{3/2}[\xi_2-\xi_1^2(1-t_{\xi_0})^{-2}]^{-3/2}(1-t_{\xi_0})^{-1/2}\nu_-(\xi_0)\frac{dy}{dz}dz
\end{aligned}
$$

where $\theta_0 = (e^{2a/m_0} - 1)^{1/2}$, $\theta_1 = (e^{2a/m} - 1)^{1/2}$, $\nu_+(\theta) = \lim_{b\to\infty} E_\theta\{e^{-(U_{\tau_b}-b)}\}$ with $U_n = \sum_{i=1}^n Y_i$, $Y_1, Y_2, \cdots$ is an i.i.d. sequence of random variables, each having the same distribution as

$$Y = -(\theta^2/1+\theta^2)X^2/2 + (\theta/1+\theta^2)x + [\theta^2/(1+\theta^2) + \log(1+\theta^2)]/2, \tag{5.2}$$

with X distributed according to the standard normal distribution, and $\tau_b = \inf\{n,\ U_n \geq b\}$.

The original forward method (see Lalley (1983)) gives the same result. (Private communication). Woodroofe (1978, 1979) contains a mistake in the general approximation formula, so his results on repeated T-tests are also incorrect. The derivation of (5.1) using Siegmund's method is simpler than that of the original forward method because it takes advantage of the invariance property of the generalized likelihood ratio statistic. The backward method also makes use of this invariance property in an implicit way. More on this point later.

By (3.5)

$$\begin{aligned} P_\xi^{(m)}\{m_0 \leq \tau < m\} \sim (1-t_{\xi_0})^{-1/2} |\Sigma(\xi_0)|^{1/2} \cdot \\ |\Sigma[\xi_1/(1-t_0), \xi_2]|^{-1/2} \exp[-(a_0 - \Lambda(\xi_0))m] \cdot \nu_-(\xi_1, \xi_1) \end{aligned} \tag{5.3}$$

where

$$\xi m^{-1} = \xi_0 = (\xi_1, \xi_1),\ \ a = a_0 m,\ \ t_{\xi_0} = \inf\{t : 0 \leq t \leq 1 - m_0/m,$$
$$(1-t)\Lambda(\xi_1/(1-t), \xi_2) = a_0\}$$

$\Sigma(\xi_1, \xi_2)$ is the covariance matrix of (W^2, W). W is normally distributed with first and second moment given by $E(W) = \xi_1$ $E(W^2) = \xi_2$. Simple calculation shows that the determinant of Σ is $|\Sigma(\xi_1, \xi_2)| = 2(\xi_2 - \xi_1^2)^3$.

$$\nu_-(\xi_1, \xi_2) = \lim_{m\to\infty} E_{\xi_0^{(2)}}^{(m)}(e^{-R_m}),\ \ R_m = (m-T)\Lambda[(\xi - S_T)/(m-T)] - a$$

$$T = \inf\{n;\ (m-n)\Lambda[(\xi - S_n)/(m-n)] > a\}.$$

By Theorem 3 of Chapter 4, R_m under $P_{\xi^{(2)}}^{(m)}$ has the same limiting distribution as the excess over the boundary by a random walk with increment $\nabla H(t_{\xi_0}\mu) \cdot (1,\ \sqrt{\xi_2}X,\ \xi_2 X^2)$ where $\mu = (1, 0, \xi_2)$ is the mean vector of $(1,\ \sqrt{\xi_0}x,\ \xi_2 x^2)$, $H(x,y,z) = (1-x)\Lambda[((\xi_1 - y)/(1-$

Chapter 5

An Example: The Repeated T-Test

Assume $X_1, X_2, \cdots$ are independent and normally distributed with unknown mean μ and variance σ^2, and that we are interested in testing $H_0 : \mu = 0$ against $H_1 : \mu \neq 0$. Let $\ell_{\mu,\sigma}(\cdot)$ denote the log likelihood of X_1. Some simple algebra shows that the (generalized) log likelihood ratio statistic is

$$n\Lambda(S_n/n) = n(\phi(S_n/n) - \phi_0(S_n^{(2)}/n) = (n/2)\log\left\{\frac{S_n^{(2)}}{S_n^{(2)}/n - (S_n^{(1)}/n)^2}\right\}$$

where $S_n = (S_n^{(1)}, S_n^{(2)}) = (\sum_{i=1}^n X_i,\ \sum_{i=1}^n X_i^2)$

$$\phi(x_1, x_2) = \sup_{\sigma,\mu} \ell_{\mu,\sigma}(x_1, x_2) = \frac{1}{2}[(x_2 - 1) - \log(x_2 - x_x^2)]$$

$$\phi_0(x_2) = \sup_{\sigma} \ell_{0,\sigma}(x_2) = [(x_2 - 1) - \log x_2]/2.$$

The repeated T-test is defined in terms of the following stopping rule $\tau = \inf\{n;\ n \geq m_0,\ n\Lambda(S_n/n) > a\}$. The test rejects the null hypothesis if and only if $\tau \leq m$. We first consider the problem of approximating the significance level of the repeated t-test. Observe that the probability $P_{0,\sigma}(m_0 \leq \tau < m)$ is independent of σ. So we may write $P_0\{m_0 \leq \tau < m\}$ for significance levels.

A variant of the forward method (see Siegmund (1985) for details) which involves taking the likelihood ratio of the maximum invariant process $(y_2, \cdots, y_n) = (x_1^{-1}x_2, x_1^{-1}x_3, \cdots, x_1^{-1}x_n)$ then mixing it by Lebesgue measure over the invariant paramter space gives us

$$P_0\{m_0 \leq \tau \leq m\} \sim 2(a/\pi)^{1/2}e^{-a}\int_{\theta_1}^{\theta_0}[\log(1+\theta^2)]^{-1/2}\nu_+(\theta)d\theta \tag{5.1}$$

Table 5.6

Powers of Group Repeated t-Test

# of observations in a group	a	m_0	m	η	analytic approximation	Monte Carlo (2000 replications)
2	3.65	8	40	0.7	0.962	0.961 ± 0.004
				0.6	0.880	0.888 ± 0.007
				0.5	0.726	0.741 ± 0.010
3	3.6	10	55	0.6	0.969	0.966 ± 0.004
				0.4	0.681	0.685 ± 0.010
4	3.6	10	70	0.5	0.949	0.940 ± 0.005
				0.3	0.527	0.518 ± 0.011
5	3.6	10	80	0.5	0.973	0.961 ± 0.004
				0.4	0.855	0.843 ± 0.008
				0.3	0.590	0.574 ± 0.011
7	3.6	15	120	0.4	0.970	0.966 ± 0.004
				0.3	0.790	0.773 ± 0.009
				0.2	0.420	0.414 ± 0.011

Appendix

Here we give some details about the numerical computation performed in Chapter 5.

To approximate $P_\eta\{m_0 \leq \tau < m, \Lambda(S_m/m) < c/m\}$ for $\eta \neq 0$ we use (5.20). For $P_0\{m_0\{\leq \tau < m, \Lambda(S_m/m) < c/m\}$ with $c < a$ we invoke (5.10). To approximate the significance levels of repeated the significance levels of the repeated significance test $P_0\{m_0 \leq \tau < m\}$ we use

$$P_0\{m_0 \leq \tau \leq m\} = P_0\{m_0 \leq \tau < m, \Lambda(S_m/m) \leq a/m\} + P_0\{\Lambda(S_m/m) > a\}. \qquad (A.1)$$

Replacing c by a in formula (5.10) gives the approximation of the first term on the right hand side of (A.1). The second term is easily computed by calculating the tail probability of the t-distribution.

To complete the approximations above, we need to compute $\nu_+(\theta)$ (the excess over the boundary by the forward process) numerically. The following proposition is useful.

Propositon 1: Let $Y_1, Y_2, \cdots$ be independent and identically distributed nonarithmetic random variables with a finite positive mean $\mu > 0$. For $b \geq 0$ define $S_n \sum_{i=1}^{n} Y_i$ and $\tau_b = \inf\{n, S_n > b\}$. Then

$$\begin{aligned}\lim_{b\to\infty} E\{\exp[-\alpha(S_{\tau_b} - b)]\} = \exp\{\pi^{-1}\Big[\int_0^\infty \alpha^2(\alpha^2+t^2)^{-1}t^{-1}[I\delta(t) - \pi/2]dt \\ - \int_0^\infty \alpha(\alpha^2+t^2)^{-1}(R\delta(t) + \log \mu t)dt\Big]\}\end{aligned} \qquad (A.2)$$

where $R\delta$ and $I\delta$ are the real and imaginary part of $\delta(t) = \log[1/(1-f(t))]$ with $f(t) = Ee^{itY_1}$, the characteristic function of Y_1.

Proof: See Woodroofe (1979).

Now it is sufficient to identify $f(t)$ for Y_1 given by

$$Y_1 = -(\theta^2/1+\theta^2)X^2/2 + (\theta/1+\theta^2)X + [\theta^2/1+\theta^2 + \log(1+\theta^2)]/2$$

with $X \sim N(0,1)$.

Some straightforward algebra shows that

$$\begin{aligned}
f(t) &= (1+\theta^2)^{1/2}(1+\theta^2+it\theta^2)^{-1/2}\exp\{(it/2)[\log(1+\theta^2)+\theta^2/(1+\theta^2)]\} \\
&\quad \cdot \exp\{(-1/2)\theta^2t^2\cdot(1+\theta^2)^{-1}(1+\theta^2+it\theta^2)^{-1}\} \\
&= \exp\{(1/2)\log\{(1+\theta^2)[(1+\theta^2)^2+t^2\theta^4]^{-1/2}\} - (t^2\theta^2/2)[(1+\theta^2)^2+t^2\theta^4]^{-1}\} \\
&\quad \cdot \exp\{(i/2)[t\theta^2(1+\theta^2)^{-1} - \tan^{-1}[t\theta^2(1+\theta^2)^{-1}] + t^3\theta^4[(1+\theta^2)^2+t^2\theta^4]^{-1} \\
&\quad \cdot (1+\theta^2)^{-1} + t\log(1+\theta^2)]\}
\end{aligned}$$

so

$$\begin{aligned}
|1-f(t)|^2 &= 1+\exp\{\log\{(1+\theta^2)[(1+\theta^2)^2+t^2\theta^4]^{-1/2}\} - t^2\theta^2[(1+\theta^2)^2+t^2\theta^4]^{-1}\} \\
&\quad - 2\exp\{(1/2)\log\{(1+\theta^2)[(1+\theta^2)^2+t^2\theta^4]^{-1/2}\} - (t^2\theta^2/2)[(1+\theta^2)^2+t^2\theta^4]^{-1}\} \\
&\quad \cos\{(1/2)[t\theta^2(1+\theta^2)^{-1} - \tan^{-1}[t\theta^2(1+\theta^2)^{-1}] + t^3\theta^4[(1+\theta^2)^2+t^2\theta^4]^{-1} \\
&\quad \cdot (1+\theta^2)^{-1} + t\log(1+\theta^2)]\}.
\end{aligned}$$

It is easy to see that

$$R\delta(t) = -\log|1-f(t)| \tag{A.3}$$

and

$$\begin{aligned}
I\delta(t) &= \sin^{-1}\{[|f(t)|/|1-f(t)|]\cdot\sin\{(1/2)[t\theta^2(1+\theta^2)^{-1} \\
&\quad - \tan^{-1}[t\theta^2(1+\theta^2)^{-1}] + t^3\theta^4[(1+\theta^2)^2+t^2\theta^4]^{-1} + t\log(1+\theta^2)]\}\}.
\end{aligned} \tag{A.4}$$

Now substitute (A.3) and (A.4) into (A.2) and perform the numerical integration to obtain $\nu_+(\theta)$.

References

Borisov, I. S. (1978). On the rate of convergence in the "conditional" invariance principle. *Theory Prob. Appl.*, **23**, 63–76.

Borovkov, A, A. and Rogozin, B. A. (1965). On the multidimensional central limit theorem. *Theory Prob. Appl.*, **10**, 55–62.

Chung, K. L. (1977). *A Course in Probability Theory*, 2nd ed., Academic Press, New York.

Feller, W. (1966). *An Introduction to Probability Theory and Its Applications*, Vol. 2, John Wiley & Sons, New York.

Hogan, M. (1984). Problems in boundary crossing for random walks, Ph.D. dissertation, Stanford University.

Lai, T. L. and Siegmund, D. (1977). A nonliner renewal theory with application to sequential analysis I, *Ann. Statist.*, **5**, 946–954.

Lalley, S. (1983). Repeated likelihood ratio test for curved exponential families, *Z. Wahrsch. Verw. Gebiete*, **62**, 293–321.

Siegmund, D. (1975). Importance sampling in the Monte Carlo study of sequential test, *Ann. Statist.*, **4**, 673–684.

Siegmund, D. (1977). Repeated significance test for a normal mean, *Biometrika*, **64**, 177–189.

Siegmund, D. (1980). Sequential χ^2 and F tests and related confidence intervals, *Biometrika*, **67**, 389–402.

Siegmund, D. (1982). Large deviations for boundary crossing probabilities, *Ann. Probab.*, **10**, 581–588.

Siegmund, D. (1985). Boundary crossing probabilities and statistical applications, submitted *Ann. Statist.*

Siegmund, D. (1985). *Sequential Analysis*, Springer-Verlag, New York-Heidelberg-Berlin.

Woodroofe, M. (1976). A renewal theorem for curved boundaries and moments of first passage times, *Ann. Probab.*, **4**, 67–80.

Woodroofe, M. (1977). Second order approximation for sequential point and interval estimation, *Ann. Statist.*, **5**, 984–995.

Woodroofe, M. (1978). Large deviations of likelihood ratio statistics with application to sequential testing, *Ann. Statist.*, **6**, 72–84.

Woodroofe, M. (1979). Repeated likelihood ratio tests, *Biometrika* **66**, 453–463.

Woodroofe, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis*, Society for Industrial and Applied Mathematics, Philadelphia.

Woodroofe, M. and Takahashi, H. (1982). Asymptotic expansions for the error probabilities of some replaced significance tests, *Ann. Statist.*, **10**, 895–908.

Zabell, S. (1980). Rates of convergence for conditional expectations, *Ann. Probab.*, **8**, 928–941.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 34	2. GOVT ACCESSION NO. AD-A159547	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Repeated Significance Tests for Exponential Families		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Inchi Hu		8. CONTRACT OR GRANT NUMBER(s) N00014-77-C-0306
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, California 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-373
11. CONTROLLING OFFICE NAME AND ADDRESS Statistics & Probability Program Code (411 (SP)) Office of Naval Research Arlington, Virginia 22217		12. REPORT DATE August 1985
		13. NUMBER OF PAGES 56
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) see reverse side		

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

In this report I considered the significance levels and powers of repeates significance tests (RST). Typically these quantities cannot be calculated exactly and some sort of approximations are required. Satisfactory approximations of significance levels of RST for exponential families have been obtained by Lalley (1983) and Woodroofe (1978). In this report another method due to Siegmund (1985) in the special case of normal observations with known variance is developed. The main advantages that are claimed for this method are two-fold. First, it can be used to approximate the power of the RST. Second, it enables one to estimate both powers and significance levels of the modified RST. The approximations are also useful in determining confidence intervals. The proof of a nonlinear renewal theorem for conditional random walks and some numerical results are also given. Additional keywords: [illegible] function; lattice (data); random walks; null hypothesis [illegible]

END

FILMED

10-85

DTIC

END

FILMED

10-85

DTIC